(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 May 2002 (10.05.2002)

PCT

(10) International Publication Number
WO 02/36783 A2

(51) International Patent Classification[7]: C12N 15/55, 9/12, C12Q 1/68, 1/44, A61K 31/713, 38/46, C12N 15/62, C07K 16/40, C12N 15/11, 9/16 // A61P 35/00

(21) International Application Number: PCT/EP01/12517

(22) International Filing Date: 30 October 2001 (30.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/244,183 31 October 2000 (31.10.2000) US
60/317,965 10 September 2001 (10.09.2001) US

(71) Applicant *(for all designated States except US)*: **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **XIAO, Yonghong** [US/US]; 75 Dana Street, #1, Cambridge, MA 02138 (US).

(74) Common Representative: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/36783 A2

(54) Title: REGULATION OF HUMAN HISTONE DEACETYLASE

(57) Abstract: Reagents which regulate human histone deacetylase and reagents which bind to human histone deacetylase gene products can play a role in preventing, ameliorating, or correcting dysfunctions or diseases including, but not limited to, cancer.

REGULATION OF HUMAN HISTONE DEACETYLASE

This application incorporates by reference co-pending applications Serial No. 60/244,183, filed October 31, 2000 and Serial No. 60/317,965, filed September 10, 2001.

TECHNICAL FIELD OF THE INVENTION

The invention relates to the area of enzyme regulation. More particularly, the invention relates to the regulation of human histone deacetylase and its regulation.

BACKGROUND OF THE INVENTION

Histone deacetylase and histone acetyltransferase together control the net level of acetylation of histones. U.S. Patent No. 6,110,697; Cress & Seto, *J. Cell. Physiol. 184*, 1- 16, 2000; Hu *et al.*, *J. Biol. Chem. 275*, 15254-64, 2000; Davie & Spencer, *J. Cell. Biochem. Suppl. 32-33*, 141-48, 1999. Inhibition of the action of histone deacetylase results in the accumulation of hyperacetylated histones, which in turn is implicated in a variety of cellular responses, including altered gene expression, cell differentiation and cell-cycle arrest. Thus, agents which regulate the activity of histone deacetylase can be useful as therapeutic agents for a wide variety of disorders.

SUMMARY OF THE INVENTION

It is an object of the invention to provide reagents and methods of regulating a human histone deacetylase. This and other objects of the invention are provided by one or more of the embodiments described below.

One embodiment of the invention is a histone deacetylase polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 48% identical to the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 48% identical to the amino acid sequence shown in SEQ ID NO: 7; and

the amino acid sequence shown in SEQ ID NO: 7.

Yet another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a histone deacetylase polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 48% identical to the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 48% identical to the amino acid sequence shown in SEQ ID NO: 7; and

the amino acid sequence shown in SEQ ID NO: 7.

Binding between the test compound and the histone deacetylase polypeptide is detected. A test compound which binds to the histone deacetylase polypeptide is

thereby identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the activity of the histone deacetylase.

Another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a polynucleotide encoding a histone deacetylase polypeptide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 6; and

the nucleotide sequence shown in SEQ ID NO: 6.

Binding of the test compound to the polynucleotide is detected. A test compound which binds to the polynucleotide is identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the amount of the histone deacetylase through interacting with the histone deacetylase mRNA.

Another embodiment of the invention is a method of screening for agents which regulate extracellular matrix degradation. A test compound is contacted with a histone deacetylase polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 48% identical to the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 48% identical to the amino acid sequence shown in SEQ ID NO: 7; and

the amino acid sequence shown in SEQ ID NO: 7.

A histone deacetylase activity of the polypeptide is detected. A test compound which increases histone deacetylase activity of the polypeptide relative to histone deacetylase activity in the absence of the test compound is thereby identified as a potential agent for increasing extracellular matrix degradation. A test compound which decreases histone deacetylase activity of the polypeptide relative to histone deacetylase activity in the absence of the test compound is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Even another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a histone deacetylase product of a polynucleotide which comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 6; and

the nucleotide sequence shown in SEQ ID NO: 6.

Binding of the test compound to the histone deacetylase product is detected. A test compound which binds to the histone deacetylase product is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Still another embodiment of the invention is a method of reducing extracellular matrix degradation. A cell is contacted with a reagent which specifically binds to a polynucleotide encoding a histone deacetylase polypeptide or the product encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 6; and

the nucleotide sequence shown in SEQ ID NO: 6.

Histone deacetylase activity in the cell is thereby decreased.

The invention thus provides a human histone deacetylase which can be used to identify test compounds which may act, for example, as activators or inhibitors at the enzyme's active site. Human histone deacetylase and fragments thereof also are useful in raising specific antibodies which can block the enzyme and effectively reduce its activity.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows the DNA-sequence encoding a histone deacetylase Polypeptide (SEQ ID NO:1).

Fig. 2 shows the amino acid sequence deduced from the DNA-sequence of Fig.1 (SEQ ID NO:2).

Fig. 3 shows the amino acid sequence of the protein identified by SwissProt Accession No. P56523 (SEQ ID NO:3).

Fig. 4 shows the DNA-sequence encoding a histone deacetylase Polypeptide (SEQ ID NO:4).

Fig. 5 shows the amino acid sequence of a histone deacetylase Polypeptide (SEQ ID NO:5).

Fig. 6 shows the BLASTP alignment of SEQ ID NO:2 with SwissProt P56523 (SEQ ID NO:3).

Fig. 7 shows the BLASTP - alignment of SEQ ID NO:2 against pdb|1C3P|1C3P-A.

Fig. 8 shows the HMMPFAM - alignment of SEQ ID NO:2 against pfam|hmm|Hist_deacetyl.

Fig. 9 shows the BLASTP-alignment of SEQ ID NO:2 against trembl|AF132608 (SEQ ID NO:5).

Fig. 10 shows the BLASTP - alignment of SEQ ID NO: 7 against swissnew|Q9UQL6|HDA5_HUMAN

Fig. 11 shows the BLASTP - alignment of SEQ ID NO: 7 against pdb|1C3R|1C3R-A

Fig. 12 shows the HMMPFAM - alignment of SEQ ID NO:7 against pfam|hmm|Hist_deacetyl

Fig. 13 shows the Alignment of SEQ ID NO:2 vs SEQ ID NO:7

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to an isolated polynucleotide encoding a histone deacetylase polypeptide and being selected from the group consisting of:

a) a polynucleotide encoding a histone deacetylase polypeptide comprising an amino acid sequence selected from the group consisting of:
amino acid sequences which are at least about 48% identical to
the amino acid sequence shown in SEQ ID NO: 2;
the amino acid sequence shown in SEQ ID NO: 2;
amino acid sequences which are at least about 48% identical to
the amino acid sequence shown in SEQ ID NO: 7; and
the amino acid sequence shown in SEQ ID NO: 7.

b) a polynucleotide comprising the sequence of SEQ ID NO: 1 or 6;

c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);

d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and

e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d).

Furthermore, it has been discovered by the present applicant that a novel histone deacetylase, particularly a human histone deacetylase, is a discovery of the present invention. Human histone deacetylase comprises the amino acid sequence shown in SEQ ID NOS:2 and 7. A coding sequence for human histone deacetylase is shown in

SEQ ID NO:1 and 6. A related EST (SEQ ID NO:4) is expressed in germinal center B cells.

Human histone deacetylase is 47 % identical over 163 amino acids to the protein identified with SwissProt Accession No. P56523 and annotated as "HISTONE DEACETYLASE CLR3" (Fig. 6). Human histone deacetylase is 77 % identical over 163 amino acids to the protein identified with SwissProt Accession No. Q9UQL6 and annotated as "HUMAN HISTONE DEACETYLASE 5" (Fig. 9).

Human histone deacetylase of the invention is expected to be useful for the same purposes as previously identified histone deacetylase enzymes. Human histone deacetylase is believed to be useful in therapeutic methods to treat disorders such as cancer. Human histone deacetylase also can be used to screen for human histone deacetylase activators and inhibitors.

Polypeptides

Human histone deacetylase polypeptides according to the invention comprise at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, or 163 contiguous amino acids selected from the amino acid sequence shown in SEQ ID NO:2 or 7 or a biologically active variant thereof, as defined below. A histone deacetylase polypeptide of the invention therefore can be a portion of a histone deacetylase protein, a full-length histone deacetylase protein, or a fusion protein comprising all or a portion of a histone deacetylase protein.

Biologically Active Variants

Human histone deacetylase polypeptide variants which are biologically active, *e.g.*, retain a histone deacetylase activity, also are histone deacetylase polypeptides. Preferably, naturally or non-naturally occurring histone deacetylase polypeptide variants have amino acid sequences which are at least about 48, 50, 55, 60, 65, or 70,

preferably about 75, 80, 85, 90, 96, 96, or 98% identical to the amino acid sequence shown in SEQ ID NO:2 or 7 or a fragment thereof. Percent identity between a putative histone deacetylase polypeptide variant and an amino acid sequence of SEQ ID NO:2 or 7 is determined using the FASTA Programm (3.34 January 2000) with an optimized, BL50 matrix (15:-5), ktup: 2, join: 36, opt: 24, gap-pen: -12/ -2, width: 16 (W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448). Overall identity can be calculated based on the alignment output.

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity of a histone deacetylase polypeptide can be found using computer programs well known in the art, such as DNASTAR software. Whether an amino acid change results in a biologically active histone deacetylase polypeptide can readily be determined by assaying for histone deacetylase activity, as described for example, in the specific examples, below.

Fusion Proteins

Fusion proteins are useful for generating antibodies against histone deacetylase polypeptide amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins which interact with portions of a histone deacetylase polypeptide. Protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage

display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A histone deacetylase polypeptide fusion protein comprises two polypeptide segments fused together by means of a peptide bond. The first polypeptide segment comprises at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, or 163 contiguous amino acids of SEQ ID NO:2 or 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 200, 300, 400, 500, 700 or 848 contiguous amino acids of SEQ ID NO:7 or of a biologically active variant, such as those described above. The first polypeptide segment also can comprise full-length histone deacetylase protein.

The second polypeptide segment can be a full-length protein or a protein fragment. Proteins commonly used in fusion protein construction include β-galactosidase, β-glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between the histone deacetylase polypeptide-encoding sequence and the heterologous protein sequence, so that the histone deacetylase polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two polypeptide segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises coding sequences selected from SEQ ID NO:1 or 6 in proper

reading frame with nucleotides encoding the second polypeptide segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Identification of Species Homologs

Species homologs of human histone deacetylase polypeptide can be obtained using histone deacetylase polypeptide polynucleotides (described below) to make suitable probes or primers for screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs which encode homologs of histone deacetylase polypeptide, and expressing the cDNAs as is known in the art.

Polynucleotides

A histone deacetylase polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for a histone deacetylase polypeptide. A coding sequence for human histone deacetylase is shown in SEQ ID NOS:1 and 6.

Degenerate nucleotide sequences encoding human histone deacetylase polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, or 98 % identical to the nucleotide sequence shown in SEQ ID NO:1 or 6 or its complement also are histone deacetylase polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species

homologs, and variants of histone deacetylase polynucleotides which encode biologically active histone deacetylase polypeptides also are histone deacetylase polynucleotides.

Identification of Polynucleotide Variants and Homologs

Variants and homologs of the histone deacetylase polynucleotides described above also are histone deacetylase polynucleotides. Typically, homologous histone deacetylase polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known histone deacetylase polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions--2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1 % SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1 % SDS, 50°C once, 30 minutes; then 2X SSC, room temperature twice, 10 minutes each--homologous sequences can be identified which contain at most about 25-30 % basepair mismatches. More preferably, homologous nucleic acid strands contain 15-25 % basepair mismatches, even more preferably 5-15 % basepair mismatches.

Species homologs of the histone deacetylase polynucleotides disclosed herein also can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of histone deacetylase polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m of a double-stranded DNA decreases by 1-1.5°C with every 1 % decrease in homology (Bonner *et al.*, *J. Mol. Biol. 81*, 123 (1973). Variants of human histone deacetylase polynucleotides or histone deacetylase polynucleotides of other species can therefore be identified by hybridizing a putative homologous histone deacetylase polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO:1 or 6 or the complement thereof to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising poly-

nucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to histone deacetylase polynucleotides or their complements following stringent hybridization and/or wash conditions also are histone deacetylase polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989, at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20°C below the calculated T_m of the hybrid under study. The T_m of a hybrid between a histone deacetylase polynucleotide having a nucleotide sequence shown in SEQ ID NO:1 or 6 or the complement thereof and a polynucleotide sequence which is at least about 50, preferably about 75, 90, 96, or 98 % identical to one of those nucleotide sequences can be calculated, for example, using the equation of Bolton and McCarthy, *Proc. Natl. Acad. Sci. U.S.A. 48*, 1390 (1962):

$T_m = 81.5°C - 16.6(\log_{10}[Na^+]) + 0.41(\%G + C) - 0.63(\%formamide) - 600/l)$,
where l = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4X SSC at 65°C, or 50 % formamide, 4X SSC at 42°C, or 0.5X SSC, 0.1 % SDS at 65°C. Highly stringent wash conditions include, for example, 0.2X SSC at 65°C.

Preparation of Polynucleotides

A histone deacetylase polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, or synthesized using an amplification technique, such as the polymerase

chain reaction (PCR), or by using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated histone deacetylase polynucleotides. For example, restriction enzymes and probes can be used to isolate polynucleotide fragments which comprises histone deacetylase nucleotide sequences. Isolated polynucleotides are in preparations which are free or at least 70, 80, or 90 % free of other molecules.

Human histone deacetylase cDNA molecules can be made with standard molecular biology techniques, using histone deacetylase mRNA as a template. Human histone deacetylase cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook *et al.* (1989). An amplification technique, such as PCR, can be used to obtain additional copies of polynucleotides of the invention, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesizes histone deacetylase polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a histone deacetylase polypeptide having, for example, an amino acid sequence shown in SEQ ID NO:2 or 7 or a biologically active variant thereof.

Extending Polynucleotides

The partial sequence disclosed herein can be used to identify the corresponding full length gene from which it was derived. The partial sequence can be nick-translated or end-labeled with ^{32}P using polynucleotide kinase using labeling methods known to those with skill in the art (BASIC METHODS IN MOLECULAR BIOLOGY, Davis *et al.*, eds., Elsevier Press, N.Y., 1986). A lambda library prepared from human tissue can be directly screened with the labeled sequences of interest or the library can be converted en masse to pBluescript (Stratagene Cloning Systems, La Jolla, Calif.

92037) to facilitate bacterial colony screening (see Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, Cold Spring Harbor Laboratory Press (1989, pg. 1.20).

Both methods are well known in the art. Briefly, filters with bacterial colonies containing the library in pBluescript or bacterial lawns containing lambda plaques are denatured, and the DNA is fixed to the filters. The filters are hybridized with the labeled probe using hybridization conditions described by Davis *et al.*, 1986. The partial sequences, cloned into lambda or pBluescript, can be used as positive controls to assess background binding and to adjust the hybridization and washing stringencies necessary for accurate clone identification. The resulting autoradiograms are compared to duplicate plates of colonies or plaques; each exposed spot corresponds to a positive colony or plaque. The colonies or plaques are selected, expanded and the DNA is isolated from the colonies for further analysis and sequencing.

Positive cDNA clones are analyzed to determine the amount of additional sequence they contain using PCR with one primer from the partial sequence and the other primer from the vector. Clones with a larger vector-insert PCR product than the original partial sequence are analyzed by restriction digestion and DNA sequencing to determine whether they contain an insert of the same size or similar as the mRNA size determined from Northern blot Analysis.

Once one or more overlapping cDNA clones are identified, the complete sequence of the clones can be determined, for example after exonuclease III digestion (McCombie *et al.*, *Methods 3*, 33-40, 1991). A series of deletion clones are generated, each of which is sequenced. The resulting overlapping sequences are assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a highly accurate final sequence.

Various PCR-based methods can be used to extend the nucleic acid sequences disclosed herein to detect upstream sequences such as promoters and regulatory elements. For example, restriction-site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, *PCR Methods Applic. 2*, 318-322, 1993). Genomic DNA is first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (Triglia *et al.*, *Nucleic Acids Res. 16*, 8186, 1988). Primers can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 22-30 nucleotides in length, to have a GC content of 50 % or more, and to anneal to the target sequence at temperatures about 68-72 °C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom *et al.*, *PCR Methods Applic. 1*, 111-119, 1991). In this method, multiple restriction enzyme digestions and ligations also can be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Another method which can be used to retrieve unknown sequences is that of Parker *et al.*, *Nucleic Acids Res. 19*, 3055-3060, 1991). Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to

walk genomic DNA (CLONTECH, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Randomly-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' non-transcribed regulatory regions.

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (*e.g.* GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the entire process from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

Obtaining Polypeptides

Human histone deacetylase polypeptides can be obtained, for example, by purification from human cells, by expression of histone deacetylase polynucleotides, or by direct chemical synthesis.

Protein Purification

Human histone deacetylase polypeptides can be purified from any cell which expresses the enzyme, including host cells which have been transfected with histone deacetylase expression constructs. A purified histone deacetylase polypeptide is separated from other compounds which normally associate with the histone deacetylase polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified histone deacetylase polypeptides is at least 80 % pure; preferably, the preparations are 90 %, 95 %, or 99 % pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis.

Expression of Polynucleotides

To express a histone deacetylase polynucleotide, the polynucleotide can be inserted into an expression vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding histone deacetylase polypeptides and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook *et al.* (1989) and in Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.

A variety of expression vector/host systems can be utilized to contain and express sequences encoding a histone deacetylase polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant

bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems infected with virus expression vectors (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those non-translated regions of the vector -- enhancers, promoters, 5' and 3' untranslated regions -- which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (*e.g.*, heat shock, RUBISCO, and storage protein genes) or from plant viruses (*e.g.*, viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a histone deacetylase polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

Bacterial and Yeast Expression Systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the histone deacetylase polypeptide. For example, when a large quantity of a histone deacetylase polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT

(Stratagene). In a BLUESCRIPT vector, a sequence encoding the histone deacetylase polypeptide can be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β-galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem. 264*, 5503-5509, 1989) or pGEX vectors (Promega, Madison, Wis.) also can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel *et al.* (1989) and Grant *et al.*, *Methods Enzymol. 153*, 516-544, 1987.

Plant and Insect Expression Systems

If plant expression vectors are used, the expression of sequences encoding histone deacetylase polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be used alone or in combination with the omega leader sequence from TMV (Takamatsu, *EMBO J. 6*, 307-311, 1987). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used (Coruzzi *et al.*, *EMBO J. 3*, 1671-1680, 1984; Broglie *et al.*, *Science 224*, 838-843, 1984; Winter *et al.*, *Results Probl. Cell Differ. 17*, 85-105, 1991). These constructs can be introduced into plant cells by direct DNA transformation or by pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (*e.g.*, Hobbs or Murray, in MCGRAW HILL YEARBOOK OF SCIENCE AND TECHNOLOGY, McGraw Hill, New York, N.Y., pp. 191-196, 1992).

An insect system also can be used to express a histone deacetylase polypeptide. For example, in one such system *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. Sequences encoding histone deacetylase polypeptides can be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of histone deacetylase polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect *S. frugiperda cells* or *Trichoplusia* larvae in which histone deacetylase polypeptides can be expressed (Engelhard *et al.*, *Proc. Nat. Acad. Sci. 91*, 3224-3227, 1994).

Mammalian Expression Systems

A number of viral-based expression systems can be used to express histone deacetylase polypeptides in mammalian host cells. For example, if an adenovirus is used as an expression vector, sequences encoding histone deacetylase polypeptides can be ligated into an adenovirus transcription/translation complex comprising the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome can be used to obtain a viable virus which is capable of expressing a histone deacetylase polypeptide in infected host cells (Logan & Shenk, *Proc. Natl. Acad. Sci. 81*, 3655-3659, 1984). If desired, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (*e.g.*, liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding histone deacetylase polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a histone deacetylase polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used (see Scharf *et al.*, *Results Probl. Cell Differ. 20*, 125-162, 1994).

Host Cells

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed histone deacetylase polypeptide in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (*e.g.*, CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express histone deacetylase polypeptides can be transformed using expression vectors which can contain viral origins

of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 1-2 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced histone deacetylase sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. See, for example, ANIMAL CELL CULTURE, R.I. Freshney, ed., 1986.

Any number of selection systems can be used to recover transformed cell lines.

These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, *Cell 11*, 223-32, 1977) and adenine phosphoribosyltransferase (Lowy *et al.*, *Cell 22*, 817-23, 1980) genes which can be employed in *tk*$^-$ or *aprt*$^-$ cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler *et al.*, *Proc. Natl. Acad. Sci. 77*, 3567-70, 1980), *npt* confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin *et al.*, *J. Mol. Biol. 150*, 1-14, 1981), and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murray, 1992, *supra*). Additional selectable genes have been described. For example, *trpB* allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci. 85*, 8047-51, 1988). Visible markers such as anthocyanins, β-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes *et al.*, *Methods Mol. Biol. 55*, 121-131, 1995).

Detecting Expression

Although the presence of marker gene expression suggests that the histone deacetylase polynucleotide is also present, its presence and expression may need to be confirmed. For example, if a sequence encoding a histone deacetylase polypeptide is inserted within a marker gene sequence, transformed cells containing sequences which encode a histone deacetylase polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a histone deacetylase polypeptide under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the histone deacetylase polynucleotide.

Alternatively, host cells which contain a histone deacetylase polynucleotide and which express a histone deacetylase polypeptide can be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein. For example, the presence of a polynucleotide sequence encoding a histone deacetylase polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a histone deacetylase polypeptide. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding a histone deacetylase polypeptide to detect transformants which contain a histone deacetylase polynucleotide.

A variety of protocols for detecting and measuring the expression of a histone deacetylase polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a histone deacetylase poly-

peptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton *et al.*, SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med. 158*, 1211-1216, 1983).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding histone deacetylase polypeptides include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding a histone deacetylase polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Expression and Purification of Polypeptides

Host cells transformed with nucleotide sequences encoding a histone deacetylase polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode histone deacetylase polypeptides can be designed to contain signal sequences which direct secretion of soluble histone deacetylase polypeptides through a prokaryotic or eukaryotic cell membrane or which direct the membrane insertion of membrane-bound histone deacetylase polypeptide.

As discussed above, other constructions can be used to join a sequence encoding a histone deacetylase polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). Inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the histone deacetylase polypeptide also can be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a histone deacetylase polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilized metal ion affinity chromatography, as described in Porath *et al.*, *Prot. Exp. Purif. 3*, 263-281, 1992), while the enterokinase cleavage site provides a means for purifying the histone deacetylase polypeptide from the fusion protein. Vectors which contain fusion proteins are disclosed in Kroll *et al.*, *DNA Cell Biol. 12*, 441-453, 1993.

Chemical Synthesis

Sequences encoding a histone deacetylase polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers *et al.*, *Nucl. Acids Res. Symp. Ser.* 215-223, 1980; Horn *et al. Nucl. Acids Res. Symp. Ser.* 225-232, 1980). Alternatively, a histone deacetylase polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques (Merrifield, *J. Am. Chem. Soc. 85*, 2149-2154, 1963; Roberge *et al.*, *Science 269*, 202-204, 1995). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer

(Perkin Elmer). Optionally, fragments of histone deacetylase polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES, WH Freeman and Co., New York, N.Y., 1983). The composition of a synthetic histone deacetylase polypeptide can be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure; *see* Creighton, *supra*). Additionally, any portion of the amino acid sequence of the histone deacetylase polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

Production of Altered Polypeptides

As will be understood by those of skill in the art, it may be advantageous to produce histone deacetylase polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter histone deacetylase polypeptide-encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the polypeptide or mRNA product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new

restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of a histone deacetylase polypeptide. "Antibody" as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, $F(ab')_2$, and Fv, which are capable of binding an epitope of a histone deacetylase polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, *e.g.*, at least 15, 25, or 50 amino acids.

An antibody which specifically binds to an epitope of a histone deacetylase polypeptide can be used therapeutically, as well as in immunochemical assays, such as Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody which specifically binds to the immunogen.

Typically, an antibody which specifically binds to a histone deacetylase polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to histone acetylase polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a histone deacetylase polypeptide from solution.

Human histone deacetylase polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, a histone deacetylase polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (*e.g.*, aluminum hydroxide), and surface active substances (*e.g.* lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used in humans, BCG (*bacilli Calmette-Guerin*) and *Corynebacterium parvum* are especially useful.

Monoclonal antibodies which specifically bind to a histone deacetylase polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler *et al.*, *Nature 256*, 495-497, 1985; Kozbor *et al.*, *J. Immunol. Methods 81*, 31-42, 1985; Cote *et al.*, *Proc. Natl. Acad. Sci. 80*, 2026-2030, 1983; Cole *et al.*, *Mol. Cell Biol. 62*, 109-120, 1984).

In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison *et al.*, *Proc. Natl. Acad. Sci. 81*, 6851-6855, 1984; Neuberger *et al.*, *Nature 312*, 604-608, 1984; Takeda *et al.*, *Nature 314*, 452-454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively,

humanized antibodies can be produced using recombinant methods, as described in GB2188638B. Antibodies which specifically bind to a histone deacetylase polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies which specifically bind to histone deacetylase polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobin libraries (Burton, *Proc. Natl. Acad. Sci. 88*, 11120-23, 1991).

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J. Cancer Prev. 5*, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol. 15*, 159-63. Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, 1994, *J. Biol. Chem. 269*, 199-206.

A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology (Verhaar *et al.*, 1995, *Int. J. Cancer 61*, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth. 165*, 81-91).

Antibodies which specifically bind to histone deacetylase polypeptides also can be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as

disclosed in the literature (Orlandi *et al.*, *Proc. Natl. Acad. Sci. 86*, 3833-3837, 1989; Winter *et al.*, *Nature 349*, 293-299, 1991).

Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in WO 93/03151. Binding proteins which are derived from immunoglobulins and which are multivalent and multispecific, such as the "diabodies" described in WO 94/13804, also can be prepared.

Antibodies according to the invention can be purified by methods well known in the art. For example, antibodies can be affinity purified by passage over a column to which a histone deacetylase polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

Antisense Oligonucleotides

Antisense oligonucleotides are nucleotide sequences which are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form complexes and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 11 nucleotides in length, but can be at least 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of histone deacetylase gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates,

alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters. *See* Brown, *Meth. Mol. Biol. 20*, 1-8, 1994; Sonveaux, *Meth. Mol. Biol. 26*, 1-72, 1994; Uhlmann *et al.*, *Chem. Rev. 90*, 543-583, 1990.

Modifications of histone deacetylase gene expression can be obtained by designing antisense oligonucleotides which will form duplexes to the control, 5', or regulatory regions of the histone deacetylase gene. Oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature (*e.g.*, Gee *et al.*, in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Futura Publishing Co., Mt. Kisco, N.Y., 1994). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful complex formation between an antisense oligonucleotide and the complementary sequence of a histone deacetylase polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to a histone deacetylase polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent histone deacetylase nucleotides, can provide sufficient targeting specificity for histone deacetylase mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can easily use the calculated melting point of an antisense-sense pair to determine the degree of mismatching which will be tolerated between a particular antisense oligonucleotide and a particular histone deacetylase polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to a histone deacetylase polynucleotide. These modifications can be internal or at one or both ends of the antisense molecule. For example, internucleoside phosphate linkages can be modified by adding cholesteryl or diamine moieties with varying numbers of carbon residues between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or a 3', 5'-substituted oligonucleotide in which the 3' hydroxyl group or the 5' phosphate group are substituted, also can be employed in a modified antisense oligonucleotide. These modified oligonucleotides can be prepared by methods well known in the art. *See, e.g.*, Agrawal *et al.*, *Trends Biotechnol. 10*, 152-158, 1992; Uhlmann *et al.*, *Chem. Rev. 90*, 543-584, 1990; Uhlmann *et al.*, *Tetrahedron. Lett. 215*, 3539-3542, 1987.

Ribozymes

Ribozymes are RNA molecules with catalytic activity. *See, e.g.*, Cech, *Science 236*, 1532-1539; 1987; Cech, *Ann. Rev. Biochem. 59*, 543-568; 1990, Cech, *Curr. Opin. Struct. Biol. 2*, 605-609; 1992, Couture & Stinchcomb, *Trends Genet. 12*, 510-515, 1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (*e.g.*, Haseloff *et al.*, U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The coding sequence of a histone deacetylase polynucleotide can be used to generate ribozymes which will specifically bind to mRNA transcribed from the histone deacetylase polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have

been developed and described in the art (*see* Haseloff *et al. Nature 334*, 585-591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.*, EP 321,201).

Specific ribozyme cleavage sites within a histone deacetylase RNA target can be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. Suitability of candidate histone deacetylase RNA targets also can be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related such that upon hybridizing to the target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease histone deacetylase expression. Alternatively, if it is desired that the cells stably retain the DNA construct, the construct can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. A ribozyme-encoding DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in Haseloff *et al.*, U.S. Patent 5,641,673, ribozymes can be engineered so that ribozyme expression will occur in response to factors which induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

Differentially Expressed Genes

Described herein are methods for the identification of genes whose products interact with human histone deacetylase. Such genes may represent genes which are differentially expressed in disorders including, but not limited to, cancer. Further, such genes may represent genes which are differentially regulated in response to manipulations relevant to the progression or treatment of such diseases. Additionally, such genes may have a temporally modulated expression, increased or decreased at different stages of tissue or organism development. A differentially expressed gene may also have its expression modulated under control versus experimental conditions. In addition, the human histone deacetylase gene or gene product may itself be tested for differential expression.

The degree to which expression differs in a normal versus a diseased state need only be large enough to be visualized via standard characterization techniques such as differential display techniques. Other such standard characterization techniques by which expression differences may be visualized include but are not limited to, quantitative RT (reverse transcriptase), PCR, and Northern analysis.

Identification of Differentially Expressed Genes

To identify differentially expressed genes total RNA or, preferably, mRNA is isolated from tissues of interest. For example, RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique which does not select against the isolation of mRNA

may be utilized for the purification of such RNA samples. See, for example, Ausubel *et al.*, ed.,, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. New York, 1987-1993. Large numbers of tissue samples may readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, U.S. Patent 4,843,155.

Transcripts within the collected RNA samples which represent RNA produced by differentially expressed genes are identified by methods well known to those of skill in the art. They include, for example, differential screening (Tedder *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 85*, 208-12, 1988), subtractive hybridization (Hedrick *et al.*, *Nature 308*, 149-53; Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 88*, 2825, 1984), and, preferably, differential display (Liang & Pardee, *Science 257*, 967-71, 1992; U.S. Patent 5,262,311).

The differential expression information may itself suggest relevant methods for the treatment of disorders involving the human histone deacetylase. For example, treatment may include a modulation of expression of the differentially expressed genes and/or the gene encoding the human histone deacetylase. The differential expression information may indicate whether the expression or activity of the differentially expressed gene or gene product or the human histone deacetylase gene or gene product are up-regulated or down- regulated.

Screening Methods

The invention provides assays for screening test compounds which bind to or modulate the activity of a histone deacetylase polypeptide or a histone deacetylase polynucleotide. A test compound preferably binds to a histone deacetylase polypeptide or polynucleotide. More preferably, a test compound decreases or increases histone acetylase activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

Test Compounds

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced recombinantly, or synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the "one-bead one-compound" library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. *See* Lam, *Anticancer Drug Des. 12*, 145, 1997.

Methods for the synthesis of molecular libraries are well known in the art (*see*, for example, DeWitt *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6909, 1993; Erb *et al. Proc. Natl. Acad. Sci. U.S.A. 91*, 11422, 1994; Zuckermann *et al.*, *J. Med. Chem. 37*, 2678, 1994; Cho *et al.*, *Science 261*, 1303, 1993; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2059, 1994; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2061; Gallop *et al.*, *J. Med. Chem. 37*, 1233, 1994). Libraries of compounds can be presented in solution (*see, e.g.*, Houghten, *BioTechniques 13*, 412-421, 1992), or on beads (Lam, *Nature 354*, 82-84, 1991), chips (Fodor, *Nature 364*, 555-556, 1993), bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 89*, 1865-1869, 1992), or phage (Scott & Smith, *Science 249*, 386-390, 1990; Devlin, *Science 249*, 404-406, 1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci. 97*, 6378-6382, 1990; Felici, *J. Mol. Biol. 222*, 301-310, 1991; and Ladner, U.S. Patent 5,223,409).

High Throughput Screening

Test compounds can be screened for the ability to bind to histone deacetylase polypeptides or polynucleotides or to affect histone deacetylase activity or histone deacetylase gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 50 to 500 µl. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the 96-well format.

Alternatively, "free format assays," or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by Jayawickreme *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 19,* 1614-18 (1994). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially released the compounds from the beads. Active compounds can be visualized as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa. (Nov. 7-10, 1995). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV-light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described by Salmon *et al.*, *Molecular Diversity 2*, 57-63 (1996). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support. When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

Binding Assays

For binding assays, the test compound is preferably a small molecule which binds to and occupies, for example, the active site of the histone deacetylase polypeptide, such that normal biological activity is prevented. Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules.

In binding assays, either the test compound or the histone deacetylase polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound which is bound to the histone deacetylase polypeptide can then be accomplished, for example, by direct counting of radioemmission, by scintillation counting, or by determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a histone deacetylase polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a histone deacetylase polypeptide. A microphysiometer (*e.g.*, Cytosensor™) is an analytical instrument

that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a histone deacetylase polypeptide (McConnell *et al.*, *Science 257*, 1906-1912, 1992).

Determining the ability of a test compound to bind to a histone deacetylase polypeptide also can be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA) (Sjolander & Urbaniczky, *Anal. Chem. 63*, 2338-2345, 1991, and Szabo *et al.*, *Curr. Opin. Struct. Biol. 5*, 699-705, 1995). BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore™). Changes in the optical phenomenon surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another aspect of the invention, a histone deacetylase polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent 5,283,317; Zervos *et al.*, *Cell 72*, 223-232, 1993; Madura *et al.*, *J. Biol. Chem. 268*, 12046-12054, 1993; Bartel *et al.*, *BioTechniques 14*, 920-924, 1993; Iwabuchi *et al.*, *Oncogene 8*, 1693-1696, 1993; and Brent W094/10300), to identify other proteins which bind to or interact with the histone deacetylase polypeptide and modulate its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct, polynucleotide encoding a histone deacetylase polypeptide can be fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct a DNA sequence that encodes an unidentified protein ("prey" or "sample") can be fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact *in vivo* to form an protein-dependent complex, the DNA-binding and activation domains of the

transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein which interacts with the histone deacetylase polypeptide.

It may be desirable to immobilize either the histone deacetylase polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or both of the interactants, as well as to accommodate automation of the assay. Thus, either the histone deacetylase polypeptide (or polynucleotide) or the test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach the enzyme polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide (or polynucleotide) or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a histone deacetylase polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, the histone deacetylase polypeptide is a fusion protein comprising a domain that allows the histone deacetylase polypeptide to be bound to a solid support. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the non-adsorbed histone deacetylase polypeptide; the mixture is then incubated under conditions conducive to complex

formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing proteins or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either a histone deacetylase polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated histone deacetylase polypeptides (or polynucleotides) or test compounds can be prepared from biotin-NHS(N-hydroxysuccinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.) and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which specifically bind to a histone deacetylase polypeptide, polynucleotide, or a test compound, but which do not interfere with a desired binding site, such as the active site of the histone deacetylase polypeptide, can be derivatized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies which specifically bind to the histone deacetylase polypeptide or test compound, enzyme-linked assays which rely on detecting an activity of the histone deacetylase polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a histone deacetylase polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a histone deacetylase polypeptide or polynucleotide can be used in a cell-based assay system. A histone deacetylase polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Binding of the

test compound to a histone deacetylase polypeptide or polynucleotide is determined as described above.

Enzyme Assays

Test compounds can be tested for the ability to increase or decrease the histone acetylase activity of a human histone deacetylase polypeptide. Histone acetylase activity can be measured, for example, as described in the specific examples, below.

Enzyme assays can be carried out after contacting either a purified histone deacetylase polypeptide, a cell membrane preparation, or an intact cell with a test compound. A test compound which decreases a histone acetylase activity of a histone deacetylase polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for decreasing histone deacetylase activity. A test compound which increases a histone acetylase activity of a human histone deacetylase polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100 % is identified as a potential therapeutic agent for increasing human histone deacetylase activity.

Gene Expression

In another embodiment, test compounds which increase or decrease histone deacetylase gene expression are identified. A histone deacetylase polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the histone deacetylase polynucleotide is determined. The level of expression of appropriate mRNA or polypeptide in the presence of the test compound is compared to the level of expression of mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a modulator of expression based on this comparison. For example, when expression of mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of the mRNA or polypeptide

expression. Alternatively, when expression of the mRNA or polypeptide is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of the mRNA or polypeptide expression.

The level of histone deacetylase mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or polypeptide. Either qualitative or quantitative methods can be used. The presence of polypeptide products of a histone deacetylase polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined *in vivo*, in a cell culture, or in an *in vitro* translation system by detecting incorporation of labeled amino acids into a histone deacetylase polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell which expresses a histone deacetylase polynucleotide can be used in a cell-based assay system. The histone deacetylase polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, such as CHO or human embryonic kidney 293 cells, can be used.

Pharmaceutical Compositions

The invention also provides pharmaceutical compositions which can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise, for example, a histone deacetylase polypeptide, histone deacetylase polynucleotide, ribozymes or antisense oligonucleotides, antibodies which specifically bind to a histone deacetylase polypeptide, or mimetics, activators, inhibitors, or inhibitors of a histone deacetylase polypeptide activity. The compositions can be administered alone or in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, bio-

compatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxy-propylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and

suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt

and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7 % mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

Therapeutic Indications and Methods

Increasing evidence from recent research suggests a connection between cancer and a deranged equilibrium of histone acetylation, which is maintained by two competing enzymatic activities, histone acetyltransferases (HATs) and histone deacetylases (HDACs). Zwiebel, *Leukemia 14*, 488-90, 2000; Melhick *et al.*, *Mol. Cell. Biol. 20*, 2075-86, 2000; Kosugi *et al.*, *Leukemia 13*, 1316-24, 1999; Wang *et al.*, *Cancer Res. 59*, 2766-69, 1999; Fenrick & Hiebert, *J. Cell. Biochem. Suppl. 30-31*, 194-202, 1998; Wang *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 95*, 10860-65, 1998.

A significant proportion of leukemias and possibly also solid tumors may have abnormalities involving HATs or HDACs at the genomic level through genetic mutations or chromosomal alterations. In these cases, altered levels of HATs or HDACs may derange the tightly regulated equilibrium of histone acetylation, which may affect the expression of a broad spectrum of cellular genes. On the other hand, HATs and HDACs may be carried to defined target promoters as cofactors of transcription factor-bound repressor or enhancer complexes and thereby carry out

unwanted enzymatic activities in the wrong place at the wrong time. We therefore propose a model for disease being associated with a deranged equilibrium of acetylation that affects histone proteins and promoter-bound transcription factors.

Human histone deacetylase, therefore, can be regulated to treat cancer. Cancer is a disease fundamentally caused by oncogenic cellular transformation. There are several hallmarks of transformed cells that distinguish them from their normal counterparts and underlie the pathophysiology of cancer. These include uncontrolled cellular proliferation, unresponsiveness to normal death-inducing signals (immortalization), increased cellular motility and invasiveness, increased ability to recruit blood supply through induction of new blood vessel formation (angiogenesis), genetic instability, and dysregulated gene expression. Various combinations of these aberrant physiologies, along with the acquisition of drug-resistance frequently lead to an intractable disease state in which organ failure and patient death ultimately ensue.

Most standard cancer therapies target cellular proliferation and rely on the differential proliferative capacities between transformed and normal cells for their efficacy. This approach is hindered by the facts that several important normal cell types are also highly proliferative and that cancer cells frequently become resistant to these agents. Thus, the therapeutic indices for traditional anti-cancer therapies rarely exceed 2.0.

The advent of genomics-driven molecular target identification has opened up the possibility of identifying new cancer-specific targets for therapeutic intervention that will provide safer, more effective treatments for cancer patients. Thus, newly discovered tumor-associated genes and their products can be tested for their role(s) in disease and used as tools to discover and develop innovative therapies. Genes playing important roles in any of the physiological processes outlined above can be characterized as cancer targets.

Genes or gene fragments identified through genomics can readily be expressed in one or more heterologous expression systems to produce functional recombinant proteins.

These proteins are characterized *in vitro* for their biochemical properties and then used as tools in high-throughput molecular screening programs to identify chemical modulators of their biochemical activities. Activators and/or inhibitors of target protein activity can be identified in this manner and subsequently tested in cellular and *in vivo* disease models for anti-cancer activity. Optimization of lead compounds with iterative testing in biological models and detailed pharmacokinetic and toxicological analyses form the basis for drug development and subsequent testing in humans.

This invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a test compound identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a modulating agent, an antisense nucleic acid molecule, a specific antibody, ribozyme, or a histone deacetylase polypeptide binding molecule) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

A reagent which affects histone deacetylase activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce histone deacetylase activity. The reagent preferably binds to an expression product of a human histone deacetylase gene. If the expression product is a protein, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells which have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about

30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung, liver, spleen, heart brain, lymph nodes, and skin.

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, more preferably about 1.0 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition and/or liposomes having a cholesterol backbone conjugated to polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a particular cell type, such as a cell-specific ligand exposed on the outer surface of the liposome.

Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods which are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1 μg to about 10 μg of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μg to about 5 μg of polynucleotides are combined with about 8 nmol

liposomes, and even more preferably about 1.0 μg of polynucleotides is combined with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis *et al. Trends in Biotechnol. 11*, 202-05 (1993); Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem. 263*, 621-24 (1988); Wu *et al.*, *J. Biol. Chem. 269*, 542-46 (1994); Zenke *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 87*, 3655-59 (1990); Wu *et al.*, *J. Biol. Chem. 266*, 338-42 (1991).

Determination of a Therapeutically Effective Dose

The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases histone deacetylase activity relative to the histone deacetylase activity which occurs in the absence of the therapeutically effective dose.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

Therapeutic efficacy and toxicity, *e.g.*, ED_{50} (the dose therapeutically effective in 50 % of the population) and LD_{50} (the dose lethal to 50 % of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD_{50}/ED_{50}.

Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active ingredient or to maintain the desired effect. Factors which can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either *ex vivo* or *in vivo* using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads,

protoplast fusion, viral infection, electroporation, "gene gun," and DEAE- or calcium phosphate-mediated transfection.

Effective *in vivo* dosages of an antibody are in the range of about 5 μg to about 50 μg/kg, about 50 μg to about 5 mg/kg, about 100 μg to about 500 μg/kg of patient body weight, and about 200 to about 250 μg /kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective *in vivo* dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg, about 1 μg to about 2 mg, about 5 μg to about 500 μg, and about 20 μg to about 100 μg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides which express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of a histone deacetylase gene or the activity of a histone deacetylase polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a histone deacetylase gene or the activity of a histone deacetylase polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to histone deacetylase-specific mRNA, quantitative RT-PCR, immunologic detection of a histone deacetylase polypeptide, or measurement of histone deacetylase activity.

In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act

synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Any of the therapeutic methods described above can be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

Diagnostic Methods

Human histone deacetylase also can be used in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases and abnormalities related to the presence of mutations in the nucleic acid sequences which encode the enzyme. For example, differences can be determined between the cDNA or genomic sequence encoding histone deacetylase in individuals afflicted with a disease and in normal individuals. If a mutation is observed in some or all of the afflicted individuals but not in normal individuals, then the mutation is likely to be the causative agent of the disease.

Sequence differences between a reference gene and a gene having mutations can be revealed by the direct DNA sequencing method. In addition, cloned DNA segments can be employed as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. For example, a sequencing primer can be used with a double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures using radiolabeled nucleotides or by automatic sequencing procedures using fluorescent tags.

Genetic testing based on DNA sequence differences can be carried out by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized, for

example, by high resolution gel electrophoresis. DNA fragments of different sequences can be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (*see, e.g.*, Myers *et al.*, *Science 230*, 1242, 1985). Sequence changes at specific locations can also be revealed by nuclease protection assays, such as RNase and S 1 protection or the chemical cleavage method (*e.g.*, Cotton *et al.*, *Proc. Natl. Acad. Sci. USA 85*, 4397-4401, 1985). Thus, the detection of a specific DNA sequence can be performed by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes and Southern blotting of genomic DNA. In addition to direct methods such as gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

Altered levels of a histone deacetylase also can be detected in various tissues. Assays used to detect levels of the receptor polypeptides in a body sample, such as blood or a tissue biopsy, derived from a host are well known to those of skill in the art and include radioimmunoassays, competitive binding assays, Western blot analysis, and ELISA assays.

All patents and patent applications cited in this disclosure are expressly incorporated herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLE 1

Detection of histone deacetylase activity

The polynucleotide of SEQ ID NO: 1 or 6 is inserted into the expression vector pCEV4 and the expression vector pCEV4-histone deacetylase polypeptide obtained is transfected into human embryonic kidney 293 cells. From these cells extracts are obtained and histone deacetylase activity is measured in an assay in a total volume of 40 µl : 400 nmol HEPES-sodium, pH 7.4, 100 pmol of the substrate [AcGly-Ala-Lys(-^{14}C-Ac)-Arg-His-Arg-Lys(^{14}C-Ac)-ValNH$_2$] (see Kervabon *et al.*, *FEBS Letters 106*, 93-96, 1979) having a specific activity of approximately 114 mCi/mmol, and the cell extract as deacetylase activity source. The amount of the cell extract is chosen such that about 20% of the substrate is consumed during the assay. The reaction is initiated by cell extract addition and allowed to proceed for 60 min at 41 degrees. At 60 min, the reaction is terminated by the addition of a 50% slurry of Amberlite.RTM. AG 50 W x 4 cation exchange resin, sodium form (200-400 mesh) in 25 mM sodium acetate buffer, pH 4.2 (200 µl). The resin binds both remaining substrate and the (partially) deacetylated peptidyl products. The quenched reaction is then incubated for at least 30 min at 25° with occasional mixing, diluted with additional 25 mM sodium acetate buffer, pH 4.2 (760 µ; final volume 1000 µ), incubated for a minimum of an additional 30 min at 25 degrees with occasional mixing, and then centrifuged at 10,000 x g for 1 min. An aliquot of the supernatant (800 µl) containing the enzymatically released ^{14} C-acetate is removed, mixed with Aquasol 2 liquid scintillation counter (LSC) cocktail (10 ml), and counted in a Beckman model LS-5801 LSC. To assure that the acetate released is due specifically to the action of histone deacetylase, a parallel control incubation is performed which contained a known histone deacetylase inhibitor [originally, 1-5 mM butyrate (see Cousens et al (1979) J. Biol. Chem. 254: 1716-1723); later, 40-1000 nM apicidin in DMSO once it had been demonstrated to be an histone deacetylase inhibitor]; the amount of radioactivity generated in the presence of inhibitor is subtracted from the value obtained in the absence of inhibitor in order to calculate histone deacetylase

dependent acetate production. It is shown that the polypeptide of SEQ ID NO: 2 or 7 respectively have a histone deacetylase activity.

EXAMPLE 2

Expression of recombinant human histone deacetylase

The *Pichia pastoris* expression vector pPICZB (Invitrogen, San Diego, CA) is used to produce large quantities of recombinant human histone deacetylase polypeptides in yeast. The histone deacetylase-encoding DNA sequence is derived from SEQ ID NO:1 or 6. Before insertion into vector pPICZB, the DNA sequence is modified by well known methods in such a way that it contains at its 5'-end an initiation codon and at its 3'-end an enterokinase cleavage site, a His6 reporter tag and a termination codon. Moreover, at both termini recognition sequences for restriction endonucleases are added and after digestion of the multiple cloning site of pPICZ B with the corresponding restriction enzymes the modified DNA sequence is ligated into pPICZB. This expression vector is designed for inducible expression in *Pichia pastoris*, driven by a yeast promoter. The resulting pPICZ/md-His6 vector is used to transform the yeast.

The yeast is cultivated under usual conditions in 5 liter shake flasks and the recombinantly produced protein isolated from the culture by affinity chromatography (Ni-NTA-Resin) in the presence of 8 M urea. The bound polypeptide is eluted with buffer, pH 3.5, and neutralized. Separation of the polypeptide from the His6 reporter tag is accomplished by site-specific proteolysis using enterokinase (Invitrogen, San Diego, CA) according to manufacturer's instructions. Purified human histone deacetylase polypeptide is obtained.

EXAMPLE 3

Identification of test compounds that bind to histone deacetylase polypeptides

Purified histone deacetylase polypeptides comprising a glutathione-S-transferase protein and absorbed onto glutathione-derivatized wells of 96-well microtiter plates are contacted with test compounds from a small molecule library at pH 7.0 in a physiological buffer solution. Human histone deacetylase polypeptides comprise the amino acid sequence shown in SEQ ID NO:2 or 7. The test compounds comprise a fluorescent tag. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to a histone deacetylase polypeptide is detected by fluorescence measurements of the contents of the wells. A test compound which increases the fluorescence in a well by at least 15 % relative to fluorescence of a well in which a test compound is not incubated is identified as a compound which binds to a histone deacetylase polypeptide.

EXAMPLE 4

Identification of a test compound which decreases histone deacetylase gene expression

A test compound is administered to a culture of human cells transfected with a histone deacetylase expression construct and incubated at 37°C for 10 to 45 minutes. A culture of the same type of cells which have not been transfected is incubated for the same time without the test compound to provide a negative control.

RNA is isolated from the two cultures as described in Chirgwin *et al.*, *Biochem. 18*, 5294-99, 1979). Northern blots are prepared using 20 to 30 µg total RNA and hybridized with a ^{32}P-labeled histone deacetylase-specific probe at 65°C in Express-hyb (CLONTECH). The probe comprises at least 11 contiguous nucleotides selected

from the complement of SEQ ID NO:1 or 6. A test compound which decreases the histone deacetylase- specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of histone deacetylase gene expression.

EXAMPLE 5

Identification of a test compound which decreases histone deacetylase activity

A test compound is administered to a culture of human cells transfected with a histone deacetylase expression construct and incubated at 37°C for 10 to 45 minutes. A culture of the same type of cells which have not been transfected is incubated for the same time without the test compound to provide a negative control. Histone deacetylase activity is measured using the methods described in Example 5.

A test compound which decreases the histone acetylase activity of the histone deacetylase relative to the histone acetylase activity in the absence of the test compound is identified as an inhibitor of histone deacetylase activity.

EXAMPLE 6

Histone Deacetylase Assays (all temperatures in °C.):

Assay 1 for Histone Deacetylase Activity and Inhibition. The standard assay is contained in a total volume of 40 µl : 400 nmol HEPES-sodium, pH 7.4, 100 pmol of the substrate [AcGly-Ala-Lys(-^{14}C-Ac)-Arg-His-Arg-Lys(^{14}C-Ac)-ValNH$_2$] (see Kervabon *et al.*, *FEBS Letters 106*, 93-96, 1979) having a specific activity of approximately 114 mCi/mmol, and a source of histone deacetylase (HDAase) activity. The amount of HDAase added is chosen such that about 20 % of the substrate is consumed during the assay. The reaction is initiated by enzyme addition and allowed to proceed for 60 min at 41 degrees. At 60 min, the reaction is terminated by the addition of a 50 % slurry of Amberlite.RTM. AG 50 W x 4 cation exchange resin, sodium form (200-400 mesh) in 25 mM sodium acetate buffer, pH 4.2 (200 µl). The resin binds both remaining substrate and the (partially) deacety-

lated peptidyl products. The quenched reaction is then incubated for at least 30 min at 25° with occasional mixing, diluted with additional 25 mM sodium acetate buffer, pH 4.2 (760 µl; final volume 1000 µl), incubated for a minimum of an additional 30 min at 25 degrees with occasional mixing, and then centrifuged at 10,000 x g for 1 min. An aliquot of the supernatant (800 µl) containing the enzymatically released ^{14}C-acetate is removed, mixed with Aquasol 2 liquid scintillation counter (LSC) cocktail (10 ml), and counted in a Beckman model LS-5801 LSC. To assure that the acetate released is due specifically to the action of HDAase, a parallel control incubation is performed which contained a known HDAase inhibitor [originally, 1-5 mM butyrate (see Cousens et al (1979) J. Biol. Chem. 254: 1716-1723); later, 40-1000 nM apicidin in DMSO once it had been demonstrated to be an HDAase inhibitor]; the amount of radioactivity generated in the presence of inhibitor is subtracted from the value obtained in the absence of inhibitor in order to calculate HDAase dependent acetate production.

For inhibition studies, the inhibitor under examination is added to the standard assay cocktail at the desired concentration in dimethyl sulfoxide (final concentration of DMSO in the reaction is kept constant at 2.5 % v/v) and the HDAase activity compared to that found in control (minus inhibitor) incubations which lacked inhibitor but contained 2.5 % v/v final DMSO.

Assay 2 for Histone Deacetylase Activity and Inhibition. The standard assay is contained in a total volume of 200 µl: 2000 nmol HEPES-sodium, pH 7.4, 11 pmol AcGly-Ala-Lys(^{3}H-Ac)-Arg-His-Arg-Lys(^{3}H-Ac)-ValNH_2 having a specific activity of approximately 3 Ci/mmol, and a source of histone deacetylase (HDAase) activity. The amount of HDAase added is chosen such that approximately 20 % of the substrate is consumed during the assay. The reaction is initiated by enzyme addition and allowed to proceed for 60 min at 41 degrees. At 60 min, the reaction is terminated by the addition of a aqueous solution containing 0.1 M acetic acid and 0.5 M hydrochloric acid (20 µl), followed by the addition of ethyl acetate (1000 µl). The quenched reaction is then vortexed for at least 15 sec at 25 degrees and then

centrifuged at 10,000 X g for 1 min. An aliquot of the ethyl acetate phase (900 µl) containing the enzymatically released ^{3}H-acetate is removed, mixed with Aquasol 2 liquid scintillation counter (LSC) cocktail (6 ml), and counted in a Beckman model LS-5801 LSC. To assure that the acetate released is due specifically to the action of HDAase, a parallel control incubation is performed which contained a known HDAase inhibitor [originally, 1-5 mM butyrate; later, 40-1000 nM apicidin in DMSO once it had been demonstrated to be an HDAase inhibitor]; the amount of radioactivity generated in the presence of inhibitor is subtracted from the value obtained in the absence of inhibitor in order to calculate HDAase dependent acetate production.

For inhibition studies, the inhibitor under examination is added to the standard assay cocktail at the desired concentration in dimethyl sulfoxide (final concentration of DMSO in the reaction is kept constant at 0.5 % v/v) and the HDAase activity compared to that found in control (minus inhibitor) incubations which lacked inhibitor but contained 0.5% v/v final DMSO.

EXAMPLE 7

Tissue-specific expression of histone deacetylase

The qualitative expression pattern of histone deacetylase in various tissues is determined by Reverse Transcription-Polymerase Chain Reaction (RT-PCR). To demonstrate that histone deacetylase is involved in cancer, expression is determined in the following tissues: adrenal gland, bone marrow, brain, cerebellum, colon, fetal brain, fetal liver, heart, kidney, liver, lung, mammary gland, pancreas, placenta, prostate, salivary gland, skeletal muscle, small intestine, spinal cord, spleen, stomach, testis, thymus, thyroid, trachea, uterus, and peripheral blood lymphocytes. Expression in the following cancer cell lines also is determined: DU-145 (prostate), NCI-H125 (lung), HT-29 (colon), COLO-205 (colon), A-549 (lung), NCI-H460 (lung), HT-116 (colon), DLD-1 (colon), MDA-MD-231 (breast), LS174T (colon),

ZF-75 (breast), MDA-MN-435 (breast), HT-1080, MCF-7 (breast), and U87. Matched pairs of malignant and normal tissue from the same patient also are tested.

Quantitative expression profiling. Quantitative expression profiling is performed by the form of quantitative PCR analysis called "kinetic analysis" firstly described in Higuchi *et al.*, *BioTechnology 10*, 413-17, 1992, and Higuchi *et al.*, *BioTechnology 11*, 1026-30, 1993. The principle is that at any given cycle within the exponential phase of PCR, the amount of product is proportional to the initial number of template copies.

If the amplification is performed in the presence of an internally quenched fluorescent oligonucleotide (TaqMan probe) complementary to the target sequence, the probe is cleaved by the 5'-3' endonuclease activity of Taq DNA polymerase and a fluorescent dye released in the medium (Holland *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 88*, 7276-80, 1991). Because the fluorescence emission will increase in direct proportion to the amount of the specific amplified product, the exponential growth phase of PCR product can be detected and used to determine the initial template concentration (Heid *et al.*, *Genome Res. 6*, 986-94, 1996, and Gibson *et al.*, *Genome Res. 6*, 995-1001, 1996).

The amplification of an endogenous control can be performed to standardize the amount of sample RNA added to a reaction. In this kind of experiment, the control of choice is the 18S ribosomal RNA. Because reporter dyes with differing emission spectra are available, the target and the endogenous control can be independently quantified in the same tube if probes labeled with different dyes are used.

All "real time PCR" measurements of fluorescence are made in the ABI Prism 7700.

RNA extraction and cDNA preparation. The total RNAs used for expression quantification are listed below along with their suppliers, if commercially available.

RNAs labeled "from autopsy" were extracted from autoptic tissues with the TRIzol reagent (Life Technologies, MD) according to the manufacturer's protocol.

Fifty µg of each RNA were treated with DNase I for 1 hour at 37°C in the following reaction mix: 0.2 U/µl RNase-free DNase I (Roche Diagnostics, Germany); 0.4 U/µl RNase inhibitor (PE Applied Biosystems, CA); 10 mM Tris-HCl pH 7.9; 10mM $MgCl_2$; 50 mM NaCl; and 1 mM DTT.

After incubation, RNA is extracted once with 1 volume of phenol:chloroform:isoamyl alcohol (24:24:1) and once with chloroform, and precipitated with 1/10 volume of 3 M NaAcetate, pH5.2, and 2 volumes of ethanol.

Fifty µg of each RNA from the autoptic tissues are DNase treated with the DNA-free kit purchased from Ambion (Ambion, TX). After resuspension and spectrophotometric quantification, each sample is reverse transcribed with the TaqMan Reverse Transcription Reagents (PE Applied Biosystems, CA) according to the manufacturer's protocol. The final concentration of RNA in the reaction mix is 200ng/µl. Reverse transcription is carried out with 2.5µM of random hexamer primers.

TaqMan quantitative analysis. Specific primers and probe are designed according to the recommendations of PE Applied Biosystems and are listed below:

forward primer: 5'-(gene specific sequence)-3'
reverse primer: 5'-(gene specific sequence)-3'
probe: 5'-(FAM) -(gene specific sequence) (TAMRA)-3'
where FAM = 6-carboxy-fluorescein
and TAMRA = 6-carboxy-tetramethyl-rhodamine.

The expected length of the PCR product is -(gene specific length)bp.

Quantification experiments are performed on 10 ng of reverse transcribed RNA from each sample. Each determination is done in triplicate.

Total cDNA content is normalized with the simultaneous quantification (multiplex PCR) of the 18S ribosomal RNA using the Pre-Developed TaqMan Assay Reagents (PDAR) Control Kit (PE Applied Biosystems, CA).

The assay reaction mix is as follows: 1X final TaqMan Universal PCR Master Mix (from 2X stock) (PE Applied Biosystems, CA); 1X PDAR control – 18S RNA (from 20X stock); 300 nM forward primer; 900 nM reverse primer; 200 nM probe; 10 ng cDNA; and water to 25 ml.

Each of the following steps are carried out once: pre PCR, 2 minutes at 50°C, and 10 at 95°C. The following steps are carried out 40 times: denaturation, 15 seconds at 95°C, annealing/extension, 1 minute at 60°C.

The experiment is performed on an ABI Prism 7700 Sequence Detector (PE Applied Biosystems, CA). At the end of the run, fluorescence data acquired during PCR are processed as described in the ABI Prism 7700 user's manual in order to achieve better background subtraction as well as signal linearity with the starting target quantity.

CLAIMS

1. An isolated polynucleotide encoding a histone deacetylase polypeptide and being selected from the group consisting of:

a) a polynucleotide encoding a histone deacetylase polypeptide comprising an amino acid sequence selected form the group consisting of:

amino acid sequences which are at least about 48% identical to the amino acid sequence shown in SEQ ID NO: 2;
the amino acid sequence shown in SEQ ID NO: 2;
amino acid sequences which are at least about 48% identical to the amino acid sequence shown in SEQ ID NO: 7; and
the amino acid sequence shown in SEQ ID NO: 7.

b) a polynucleotide comprising the sequence of SEQ ID NO: 1 or 6;

c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);

d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and

e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a to (d).

2. An expression vector containing any polynucleotide of claim 1.

3. A host cell containing the expression vector of claim 2.

4. A substantially purified histone deacetylase polypeptide encoded by a polynucleotide of claim 1.

5. A method for producing a histone deacetylase polypeptide, wherein the method comprises the following steps:

a) culturing the host cell of claim 3 under conditions suitable for the expression of the histone deacetylase polypeptide; and

b) recovering the histone deacetylase polypeptide from the host cell culture.

6. A method for detection of a polynucleotide encoding a histone deacetylase polypeptide in a biological sample comprising the following steps:

a) hybridizing any polynucleotide of claim 1 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and

b) detecting said hybridization complex.

7. The method of claim 6, wherein before hybridization, the nucleic acid material of the biological sample is amplified.

8. A method for the detection of a polynucleotide of claim 1 or a histone deacetylase polypeptide of claim 4 comprising the steps of:

contacting a biological sample with a reagent which specifically interacts with the polynucleotide or the histone deacetylase polypeptide.

9. A diagnostic kit for conducting the method of any one of claims 6 to 8.

10. A method of screening for agents which decrease the activity of a histone deacetylase, comprising the steps of:

contacting a test compound with any histone deacetylase polypeptide encoded by any polynucleotide of claim1;

detecting binding of the test compound to the histone deacetylase polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential therapeutic agent for decreasing the activity of a histone deacetylase.

11. A method of screening for agents which regulate the activity of a histone deacetylase, comprising the steps of:

contacting a test compound with a histone deacetylase polypeptide encoded by any polynucleotide of claim 1; and

detecting a histone deacetylase activity of the polypeptide, wherein a test compound which increases the histone deacetylase activity is identified as a potential therapeutic agent for increasing the activity of the histone deacetylase, and wherein a test compound which decreases the histone deacetylase activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity of the histone deacetylase.

12. A method of screening for agents which decrease the activity of a histone deacetylase, comprising the steps of:

contacting a test compound with any polynucleotide of claim 1 and detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of histone deacetylase.

13. A method of reducing the activity of histone deacetylase, comprising the steps of:

contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any histone deacetylase polypeptide of claim 4, whereby the activity of histone deacetylase is reduced.

14. A reagent that modulates the activity of a histone deacetylase polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claim 10 to 12.

15. A pharmaceutical composition, comprising:

the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.

16. Use of the expression vector of claim 2 or the reagent of claim 14 in the preparation of a medicament for modulating the activity of a histone deacetylase in a disease.

17. Use of claim 16 wherein the disease is cancer.

18. A cDNA encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or 7.

19. The cDNA of claim 18 which comprises SEQ ID NO:1 or 6.

20. The cDNA of claim 18 which consists of SEQ ID NO:1 or 6.

21. An expression vector comprising a polynucleotide which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or 7.

22. The expression vector of claim 21 wherein the polynucleotide consists of SEQ ID NO:1 or 6.

23. A host cell comprising an expression vector which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or 7.

24. The host cell of claim 23 wherein the polynucleotide consists of SEQ ID NO:1 or 6.

25. A purified polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or 7.

26. The purified polypeptide of claim 25 which consists of the amino acid sequence shown in SEQ ID NO:2 or 7.

27. A fusion protein comprising a polypeptide having the amino acid sequence shown in SEQ ID NO:2 or 7.

28. A method of producing a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or 7, comprising the steps of:

culturing a host cell comprising an expression vector which encodes the polypeptide under conditions whereby the polypeptide is expressed; and isolating the polypeptide.

29. The method of claim 28 wherein the expression vector comprises SEQ ID NO:1 or 6.

30. A method of detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or 7, comprising the steps of:

hybridizing a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO:1 or 6 to nucleic acid material of a biological sample, thereby forming a hybridization complex; and detecting the hybridization complex.

31. The method of claim 30 further comprising the step of amplifying the nucleic acid material before the step of hybridizing.

32. A kit for detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or 7, comprising:

a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO:1 or 6; and instructions for the method of claim 30.

33. A method of detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or 7, comprising the steps of:

contacting a biological sample with a reagent that specifically binds to the polypeptide to form a reagent-polypeptide complex; and detecting the reagent-polypeptide complex.

34. The method of claim 33 wherein the reagent is an antibody.

35. A kit for detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or 7, comprising:

an antibody which specifically binds to the polypeptide; and instructions for the method of claim 33.

36. A method of screening for agents which can modulate the activity of a human histone deacetylase, comprising the steps of:

contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 48% identical to the amino acid sequence shown in SEQ ID NO:2 or 7 and (2) the amino acid sequence shown in SEQ ID NO:2 or 7; and

detecting binding of the test compound to the polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential agent for regulating activity of the human histone deacetylase.

37. The method of claim 36 wherein the step of contacting is in a cell.

38. The method of claim 36 wherein the cell is *in vitro*.

39. The method of claim 36 wherein the step of contacting is in a cell-free system.

40. The method of claim 36 wherein the polypeptide comprises a detectable label.

41. The method of claim 36 wherein the test compound comprises a detectable label.

42. The method of claim 36 wherein the test compound displaces a labeled ligand which is bound to the polypeptide.

43. The method of claim 36 wherein the polypeptide is bound to a solid support.

44. The method of claim 36 wherein the test compound is bound to a solid support.

45. A method of screening for agents which modulate an activity of a human histone deacetylase, comprising the steps of:

contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 48% identical to the amino acid sequence shown in SEQ ID NO:2 or 7 and (2) the amino acid sequence shown in SEQ ID NO:2 or 7; and

detecting an activity of the polypeptide, wherein a test compound which increases the activity of the polypeptide is identified as a potential agent for increasing the activity of the human histone deacetylase, and wherein a test compound which decreases the activity of the polypeptide is identified as a potential agent for decreasing the activity of the human histone deacetylase.

46. The method of claim 45 wherein the step of contacting is in a cell.

47. The method of claim 45 wherein the cell is *in vitro*.

48. The method of claim 45 wherein the step of contacting is in a cell-free system.

49. A method of screening for agents which modulate an activity of a human histone deacetylase, comprising the steps of:

contacting a test compound with a product encoded by a polynucleotide which comprises the nucleotide sequence shown in SEQ ID NO:1 or 6; and

detecting binding of the test compound to the product, wherein a test compound which binds to the product is identified as a potential agent for regulating the activity of the human histone deacetylase.

50. The method of claim 49 wherein the product is a polypeptide.

51. The method of claim 49 wherein the product is RNA.

52. A method of reducing activity of a human histone deacetylase, comprising the step of:

contacting a cell with a reagent which specifically binds to a product encoded by a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO:1 or 6, whereby the activity of a human histone deacetylase is reduced.

53. The method of claim 52 wherein the product is a polypeptide.

54. The method of claim 53 wherein the reagent is an antibody.

55. The method of claim 52 wherein the product is RNA.

56. The method of claim 55 wherein the reagent is an antisense oligonucleotide.

57. The method of claim 56 wherein the reagent is a ribozyme.

58. The method of claim 52 wherein the cell is *in vitro*.

59. The method of claim 52 wherein the cell is *in vivo*.

60. A pharmaceutical composition, comprising:

a reagent which specifically binds to a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or 7; and a pharmaceutically acceptable carrier.

61. The pharmaceutical composition of claim 60 wherein the reagent is an antibody.

62. A pharmaceutical composition, comprising:

a reagent which specifically binds to a product of a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO:1 or 6; and a pharmaceutically acceptable carrier.

63. The pharmaceutical composition of claim 62 wherein the reagent is a ribozyme.

64. The pharmaceutical composition of claim 62 wherein the reagent is an antisense oligonucleotide.

65. The pharmaceutical composition of claim 62 wherein the reagent is an antibody.

66. A pharmaceutical composition, comprising:

an expression vector encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NO:2 or 7; and a pharmaceutically acceptable carrier.

67. The pharmaceutical composition of claim 66 wherein the expression vector comprises SEQ ID NO:1 or 6.

68. A method of treating a histone deacetylase dysfunction related disease, wherein the disease is cancer comprising the step of:

administering to a patient in need thereof a therapeutically effective dose of a reagent that modulates a function of a human histone deacetylase, whereby symptoms of the histone deacetylase dysfunction related disease are ameliorated.

69. The method of claim 68 wherein the reagent is identified by the method of claim 36.

70. The method of claim 68 wherein the reagent is identified by the method of claim 45.

71. The method of claim 68 wherein the reagent is identified by the method of claim 49.

Fig. 1

gtggacagtg acaccatttg gaatgagcta cactcgtccg
gtgctgcacg catggctgtt ggctgtgtca tcgagctggc
ttccaaagtg gcctcaggag agctgaagaa tgggtttgct
gttgtgaggc cccctggcca tcacgctgaa gaatccacag
ccatggggtt ctgctttttt aattcagttg caattaccgc
caaatacttg agagaccaac taaatataag caagatattg
attgtagatc tggatgttca ccatggaaac ggtacccagc
aggcctttta tgctgacccc agcatcctgt acatttcact
ccatcgctat gatgaaggga acttttccc tggcagtgga
gccccaaatg aggttggaac aggccttgga gaagggtaca
atataaatat tgcctggaca ggtggccttg atcctcccat
gggagatgtt gagtaccttg aagcattcag gttggtactt ctttctctc

Fig. 2

VDSDTIWNEL HSSGAARMAV GCVIELASKV ASGELKNGFA
VVRPPGHHAE ESTAMGFCFF NSVAITAKYL RDQLNISKIL
IVDLDVHHGN GTQQAFYADP SILYISLHRY DEGNFFPGSG
APNEVGTGLG EGYNINIAWT GGLDPPMGDV EYLEAFRLVL LSL

Fig. 3

MLASNSDGAS TSVKPSDDAV NTVTPWSILL TNNKPMSGSE
NTLNNESHEM SQILKKSGLC YDPRMRFHAT LSEVDDHPED
PRRVLRVFEA IKKAGYVSNV PSPSDVFLRI PAREATLEEL
LQVHSQEMYD RVTNTEKMSH EDLANLEKIS DSLYYNNESA
FCARLACGSA IETCTAVVTG QVKNAFAVVR PPGHHAEPHK
PGGFCLFNNV SVTARSMLQR FPDKIKRVLI VDWDIHHGNG
TQMAFYDDPN VLYVSLHRYE NGRFYPGTNY GCAENCGEGP
GLGRTVNIPW SCAGMGDGDY IYAFQRVVMP VAYEFDPDLV
IVSCGFDAAA GDHIGQFLLT PAAYAHMTQM LMGLADGKVF
ISLEGGYNLD SISTSALAVA QSLLGIPPGR LHTTYACPQA
VATINHVTKI QSQYWRCMRP KHFDANPKDA HVDRLHDVIR
TYQAKKLFED WKITNMPILR DSVSNVFNNQ VLCSSNFFQK
DNLLVIVHES PRVLGNGTSE TNVLNLNDSL LVDPVSLYVE
WAMQQDWGLI DINIPEVVTD GENAPVDILS EVKELCLYVW
DNYVELSISK NIFFIGGGKA VHGLVNLASS RNVSDRVKCM
VNFLGTEPLV GLKTASEEDL PTWYYRHSLV FVSSSNECWK
KAKRAKRRYG RLMQSEHTET SDMMEQHYRA VTQYLLHLLQ KARPTSQ

Fig. 4

GGCCTTGGAGAAGGGTACAATATAAATATTGCCTGGACAGGTGGCCTTGATCC
TCCCATGGGAGATGTTGAGTACCTTGAAGCATTCAGGACCATCGTGAAGCCTG
TGGCAAAGAGTTTGATCCAGACATGGTCTTAGTATCTGCTGGATTTGATGCAT
TGGAAGGCCACACCCCTCCTCTAGGAGGGTACAAAGTGACGGCAAAATAAACT
CCTGTGCTGGAGGTACAACAGTTTGGAAGTATACTTGGGGAAAGAGAAAACAC
AAGATGGAAGGAAGATCTCTCTTTTCACATCGGGAGCAC

Fig. 5

MNSPNESDGMSGREPSLEILPRTSLHSIPVTVEVKPVLPRAMPSSMGGGGGGS
PSPVELRGALVGSVDPTLREQQLQQELLALKQQQQLQKQLLFAEFQKQHDHLT
RQHEVQLQKHLKQQQEMLAAKQQQEMLAAKRQQELEQQRQREQQRQEELEKQR
LEQQLLILRNKEKSKESAIASTEVKLRLQEFLLSKSKEPTPGGLNHSLPQHPK
CWGAHHASLDQSSPPQSGPPGTPPSYKLPLPGPYDSRDDFPLRKTASEPNLKV
RSRLKQKVAERRSSPLLRRKDGTVISTFKKRAVEITGAGPGASSVCNSAPGSG
PSSPNSSHSTIAENGFTGSVPNIPTEMLPQHRALPLDSSPNQFSLYTSPSLPN
ISLGLQATVTVTNSHLTASPKLSTQQEAERQALQSLRQGGTLTGKFMSTSSIP
GCLLGVALEGDGSPHGHASLLQHVLLLEQARQQSTLIAVPLHGQSPLVTGERV
ATSMRTVGKLPRHRPLSRTQSSPLPQSPQALQQLVMQQQHQQFLEKQKQQQLQ
LGKILTKTGELPRQPTTHPEETEEELTEQQEVLLGEGALTMPREGSTESESTQ
EDLEEEDEEEDGEEEEDCIQVKDEEGESGAEEGPDLEEPGAGYKKLFSDAQPL
QPLQVYQAPLSLATVPHQALGRTQSSPAAPGGMKSPPDQPVKHLFTTGVVYDT
FMLKHQCMCGNTHVHPEHAGRIQSIWSRLQETGLLSKCERIRGRKATLDEIQT
VHSEYHTLLYGTSPLNRQKLDSKKLLGPISQKMYAVLPCGGIGVDSDTVWNEM
HSSSAVRMAVGCLLELAFKVAAGELKNGFAIIRPPGHHAEESTAMGFCFFNSV
AITAKLLQQKLNVGKVLIVDWDIHHGNGTQQAFYNDPSVLYISLHRYDNGNFF
PGSGAPEEVGGGPGVGYNVNVAWTGGVDPPIGDVEYLTAFRTVVMPIAHEFSP
DVVLVSAGFDAVEGHLSPLGGYSVTARCFGHLTRQLMTLAGGRVVLALEGGHD
LTAICDASEACVSALLSVELQPLDEAVLQQKPNINAVATLEKVIEIQSKHWSC
VQKFAAGLGRSLREAQAGETEEAETVSAMALLSVGAEQAQAAAAAREHSPRPAE
EPMEQEPAL

Fig. 6

BLASTP - alignment of SEQ ID NO:2 against swiss|P56523|HDA1_SCHPO

This hit is scoring at : 2e-38 (expectation value)
Alignment length (overlap) : 163
Identities : 47 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb

```
Q:      3    SDTIWNELHSSGAARMAVGCVIELASKVASGELKNGFAVVRPPGHHAEESTAMGFCFFNS
             SD::: . .S: .AR:A.G..IE..:.V.:G::KN.FAVVRPPGHHAE.... GFC.FN:
H:    150    SDSLYYNNESAFCARLACGSAIETCTAVVTGQVKNAFAVVRPPGHHAEPHKPGGFCLFNN

        VAITAKYLRDQL--NISKILIVDLDVHHGNGTQQAFYADPSILYISLHRYDEGNFFPGS-
        V::TA: : .:. .I.::LIVD D:HHGNGTQ.AFY DP::LY:SLHRY:.G.F:PG:
        VSVTARSMLQRFPDKIKRVLIVDWDIHHGNGTQMAFYDDPNVLYVSLHRYENGRFYPGTN

        -GAPNEVGTGLGEGYNINIAWTGGLDPPMGDVEYLEAFRLVLL    161
         G.....G.G G G .:NI.W: .    MGD :Y: AF: V::
        YGCAENCGEGPGLGRTVNIPWSCA---GMGDGDYIYAFQRVVM    309
```

Fig. 7

BLASTP - alignment of SEQ ID NO:2 against pdb|1C3P|1C3P-A

hdlp (histone deacetylase-like protein)
This hit is scoring at : 1e-16 (expectation value)
Alignment length (overlap) : 156
Identities : 35 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb

```
Q:      9     ELHSSGAARMAVGCVIELASKVASG-------ELKNGFAVVRPPG--HHAEESTAMGFCF
              E :: G. . .V. .: ..S.:A:G       E. .G ....P.G  HHA :S.A GFC:
H:     83     EKYNIGGYENPVSYAMFTGSSLATGSTVQAIEEFLKGNVAFNPAGGMHHAFKSRANGFCY

              FNSVAITAKYLRDQLNISKILIVDLDVHHGNGTQQAFYADPSILYISLHRYDEGNFFPGS
              .N: A: .:YLR.: ...:IL.:DLD.HH :G.Q:AFY ...:..:SLH: .E  F .
              INNPAVGIEYLRKK-GFKRILYIDLDAHHCDGVQEAFYDTDQVFVLSLHQSPEYAFPFEK

              GAPNEVGTGLGEGYNINIAWTGGLDPPMGDVEYLEA    155
              G  .E:G.G G:GYN:NI. . GL:    D E:L A
              GFLEEIGEGKGKGYNLNIPLPKGLN----DNEFLFA    233
```

Fig. 8

HMMPFAM - alignment of SEQ ID NO:2 against pfam|hmm|Hist_deacetyl

Histone deacetylase family
This hit is scoring at : 4.3; Expect = 1.5e-9
Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:      1 ----------------------------------------------VD--SDTIWNEL--
                                                         V.  .... .EL
H:      1 gyvydpevlnheckisygatHpenpeRlrlihelLleygllkkmeivtnprkatdeelll

          -HSS-------------------G---------------------------AARMAVGCV
           HS.                   G                           AAR:AVG
          vHsedyvefleslsktnleelekgtdkilleielkyfnkgdDtpvfaglyeaarlavGgs

          IELASKVASGELKNGFAVVRPPGHHAEESTAMGFCFFNSVAITAKYLRDQL--NISKILI
          :ELA.::..GEL.N.F  .  PGHHA:::.A.GFC:FN:VAI..KYL .:.   :.::LI
          leladrllegeldnafnwagGPgHHAkkgeasGFCyfNNvAiAikyllkkyPayvkRVli

          VDLDVHHGNGTQQAFYADPSILYISLHRYDEGNFFPGSGAPNEVGTGLGEGYNINIAWTg
          :D.DVHHG:GTQ:.FY D..:L :S.H:Y.:G.FFPG:G  .E:G.G G:GY.:NI. .
          iDfDvHHGDGTQeiFydddrVltvSfHkygkGefFPGtGditeiGkgkGkgytlNiPLn.

          glDPPMGDVEYLEAFRLVLL-----------------------------------------
            :. ..D .YL.AF: V:
          ..edgtdDesYlsafkhviepvleqFkPdaivisaGfDalygDptqLgsfnLtiegygem

          -------S-----L-------------------------------     163
                 :
          vrflkslaqkhcdgpllvvlEGGYtlraianvarcwialtggllg     342
```

Fig. 9

BLASTP - alignment of SEQ ID NO:2 against swissnew|Q9UQL6|HDA5_HUMAN
HISTONE DEACETYLASE 5 (HD5) (ANTIGEN NY-CO-9).//:trembl|AF132608|AF132608_1 product: "histone deacetylase 5"; Homo sapiens histone deacetylase 5 mRNA, complete cds. //:gp|AF132608|4754909 product: "histone deacetylase 5"; Homo sapiens histone deacetylase 5 mRNA, complete cds.

- This hit is scoring at : 1e-76 (expectation value)
- Alignment length (overlap) : 163
- Identities : 77 %
- Scoring matrix : BLOSUM62 (used to infer consensus pattern)
- Database searched : nrdb_1_;

```
Q:       1 VDSDTIWNELHSSGAARMAVGCVIELASKVASGELKNGFAVVRPPGHHAEESTAMGFCFF
            VDSDT:WNE:HSS.A.RMAVGC::ELA KVA:GELKNGFA::RPPGHHAEESTAMGFCFF
H:     786 VDSDTVWNEMHSSSAVRMAVGCLLELAFKVAAGELKNGFAIIRPPGHHAEESTAMGFCFF

            NSVAITAKYLRDQLNISKILIVDLDVHHGNGTQQAFYADPSILYISLHRYDEGNFFPGSG
            NSVAITAK.L:.:LN:.K:LIVD D:HHGNGTQQAFY DPS:LYISLHRYD.GNFFPGSG
            NSVAITAKLLQQKLNVGKVLIVDWDIHHGNGTQQAFYNDPSVLYISLHRYDNGNFFPGSG

            APNEVGTGLGEGYNINIAWTGGLDPPMGDVEYLEAFRLVLLSL     163
            AP.EVG G G GYN:N:AWTGG:DPP:GDVEYL.AFR.V::.:
            APEEVGGGPGVGYNVNVAWTGGVDPPIGDVEYLTAFRTVVMPI     948
```

Fig. 10

BLASTP - alignment of SEQ ID NO: 7 against swissnew|Q9UQL6|HDA5_HUMAN
Histone deacetylase 5 (HD5) (Antigen NY-CO-9).//:trembl|AF132608|AF132608_1 product: "histone deacetylase 5"; Homo sapiens histone deacetylase 5 mRNA, complete cds. //:gp|AF132608|4754909 product: "histone deacetylase 5"; Homo sapiens histone deacetylase 5 mRNA, complete cds.
This hit is scoring at : 0.0 (expectation value)
Alignment length (overlap) : 880
Identities : 56 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb_1_;

```
Q:       1 ASEPNLKVRSRLKQKVAERRSSPLLRRKDGNVVTSFKKRMFEVT-----EXXVXXXXPGX
           ASEPNLKVRSRLKQKVAERRSSPLLRRKDG.V:::FKKR..E:T     ...V ...PG.
H:     258 ASEPNLKVRSRLKQKVAERRSSPLLRRKDGTVISTFKKRAVEITGAGPGASSVCNSAPGS

           GPXXPNNGPTGXVTENETXVLPPTPHAEQMVXQQRILIHEDXMNLLXLYTXPXLPNITLG
           GP..PN:. : ..... T. :P .P  .:M: Q.R.L  :.. N ..LYT.P.LPNI:LG
           GPSSPNSSHSTIAENGFTGSVPNIP--TEMLPQHRALPLDSSPNQFSLYTSPSLPNISLG

           LPAVPX----QLNAXNXLKEKQKCETQ---TLRQGVPLPGQYGGXIPAXXXHPHVTLEGK
           L.A. .     .L.A. .L..:Q:.E.Q   :LRQG .L.G:: .. .. .    V.LEG.
           LQATVTVTNSHLTASPKLSTQQEAERQALQSLRQGGTLTGKFMSTSSIPGCLLGVALEGD

           PPNXXHQALLQHLLLKEQMRQQKLLVAGGVPLHPQSPLATKERISPGIRGTHKLPRHRPL
            . ..H.:LLQH:LL EQ.RQQ..L:A  VPLH QSPL.T ER::..:R . KLPRHRPL
           GSPHGHASLLQHVLLLEQARQQSTLIA--VPLHGQSPLVTGERVATSMRTVGKLPRHRPL

           NRTQSAPLPQS--TLAQLVIXXXHXXFLEKXKXYXXXIHMNKLLSKSIEXLKXPGSHLEE
           :RTQS:PLPQS  .L.QLV:...H..FLEK.K...  :.:.K:L:K: E. :.P :H EE
           SRTQSSPLPQSPQALQQLVMQQQHQQFLEKQKQQQ--LQLGKILTKTGELPRQPTTHPEE

           AEEELX-------GDXAMXEDRAPSS-GNSTRSDSSACVDDTLGXV--GAVKVKEEPVDS
           .EEEL.       G:.A:. .R. S: ..ST:.D ..  ::. G.   ..::VK:E. :S
           TEEELTEQQEVLLGEGALTMPREGSTESESTQEDLEEEDEEEDGEEEEDCIQVKDEEGES
```

Fig. 10 (continued)

```
--DEDAXIXEMESGEXAAFMXXPFLEPTHTRALSVRQAPLAAVGMDGLEKHRLVSRTHSS
  :E...:.E  :G ...F.... L:P     L.V QAPL:.. :.    H:.:.RT.SS
GAEEGPDLEEPGAGYKKLFSDAQPLQP-----LQVYQAPLSLATVP----HQALGRTQSS

PAASVLPHPAMDRPLQPGSATGIAYDPLMLKHQCVCGNSTTHPEHAGRIQSIWSRLQETG
PAA.   ... D:P::   .TG:.YD..MLKHQC:CGN: .HPEHAGRIQSIWSRLQETG
PAAPGGMKSPPDQPVKHLFTTGVVYDTFMLKHQCMCGNTHVHPEHAGRIQSIWSRLQETG

LLNKCERIQGRKASLEEIQLVHSEHHSLLYGTNPLDGQKLDPRILLGDDSQKFFSSLPCG
LL:KCERI:GRKA:L:EIQ.VHSE:H:LLYGT:PL: QKLD.: LLG. SQK.:: LPCG
LLSKCERIRGRKATLDEIQTVHSEYHTLLYGTSPLNRQKLDSKKLLGPISQKMYAVLPCG

GLGVDSDTIWNELHSSGAARMAVGCVIELASKVASGELKNGFAVVRPPGHHAEESTAMGF
G:GVDSDT:WNE:HSS.A.RMAVGC::ELA KVA:GELKNGFA::RPPGHHAEESTAMGF
GIGVDSDTVWNEMHSSSAVRMAVGCLLELAFKVAAGELKNGFAIIRPPGHHAEESTAMGF

CFFNSVAITAKYLRDQLNISKILIVDLDVHHGNGTQQAFYADPSILYISLHRYDEGNFFP
CFFNSVAITAK.L:.:LN:.K:LIVD D:HHGNGTQQAFY DPS:LYISLHRYD.GNFFP
CFFNSVAITAKLLQQKLNVGKVLIVDWDIHHGNGTQQAFYNDPSVLYISLHRYDNGNFFP

GSGAPNEVGTGLGEGYNINIAWTGGLDPPMGDVEYLEAFRTIVKPVAKEFDPDMVLVSAG
GSGAP.EVG G G GYN:N:AWTGG:DPP:GDVEYL.AFRT:V.P:A.EF.PD:VLVSAG
GSGAPEEVGGGPGVGYNVNVAWTGGVDPPIGDVEYLTAFRTVVMPIAHEFSPDVVLVSAG

FDALEGHTPPLGGYKVTAKCFGHLTKQLMTLADGRVVLALEGGHDLTAICDASEACVNAL
FDA:EGH..PLGGY.VTA:CFGHLT:QLMTLA.GRVVLALEGGHDLTAICDASEACV:AL
FDAVEGHLSPLGGYSVTARCFGHLTRQLMTLAGGRVVLALEGGHDLTAICDASEACVSAL

LGNELEPLAEDILHQSPNMNAVISLQKIIEIQSKYWKSVRMVAVPRGCALAGAQL--QEE
L. EL:PL E :L.Q.PN:NAV.:L:K:IEIQSK:W..V:..A.  G :L. AQ.  .EE
LSVELQPLDEAVLQQKPNINAVATLEKVIEIQSKHWSCVQKFAAGLGRSLREAQAGETEE

TETVSALASLTVDVEQPFA----QEDSRTAGEPMEEEPAL     848
.ETVSA:A L:V..EQ. A    :...R.A EPME:EPAL
AETVSAMALLSVGAEQAQAAAAREHSPRPAEEPMEQEPAL    1122
```

Fig. 11

BLASTP - alignment of SEQ ID NO: 7 against pdb|1C3R|1C3R-A
hdlp (histone deacetylase-like protein)Mutant//:pdb|1C3S|1C3S-A
hdlp (histone deacetylase-like protein)Mutant

This hit is scoring at : 7e-21 (expectation value)
Alignment length (overlap) : 272
Identities : 29 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb_1_;

```
Q:     436 HPEHAGRIQSIWSRLQETGLLNKCERIQGRKASLEEIQLVHSEHHSLLYGTNPLDGQKLD

           HP .. R:. : .  ....L::: E I:.R.A: EE: L.H:E.:

H:      20 HPLKIPRVSLLLRFKDAMNLIDEKELIKSRPATKEELLLFHTEDYI--------------

           PRILLGDDSQKFFSSLPCGGLGVDSDTIWNELHSSGAARMAVGCVIELASKVASG-----
            .:: .: SQ    S:P G.          .E :: G. . .V. .: ..S.:A:G
           NTLMEAERSQ----SVPKGA---------REKYNIGGYENPVSYAMFTGSSLATGSTVQA

           --ELKNGFAVVRPPG--HHAEESTAMGFCFFNSVAITAKYLRDQLNISKILIVDLDVHHG
             E. .G ....P.G  HHA :S.A GFC:.N: A: .:YLR.: ...:IL.:DLD.HH
           IEEFLKGNVAFNPAGGMHHAFKSRANGFCYINNPAVGIEYLRKK-GFKRILYIDLDAHHC

           NGTQQAFYADPSILYISLHRYDEGNFFPGSGAPNEVGTGLGEGYNINIAWTGGLDPPMGD
           :G.Q:AFY ...:..:SLH: .E  F   .G  .E:G.G G:GYN:NI. . GL:    D
           DGVQEAFYDTDQVFVLSLHQSPEYAFPFEKGFLEEIGEGKGKGYNLNIPLPKGLN----D

           VEYLEAFRTIVKPVAKEFDPDMVLVSAGFDAL     698
            E:L A... :: V.: F:P::.L:..G D.L
           NEFLFALEKSLEIVKEVFEPEVYLLQLGTDPL     259
```

Fig. 12

HMMPFAM - alignment of SEQ ID NO:7 against pfam|hmm|Hist_deacetyl

```
Histone deacetylase family
This hit is scoring at : 390.3
Scoring matrix : BLOSUM62 (used to infer consensus pattern)

Q:     413 SATGIAYDPLMLKHQCVCGNSTTHPEHAGRIQSIWSRLQETGLLNKCERIQ----GRKAS
           ...G..YDP :L.H:C .. ..THPE:. R:: I . L E GLL.K.E :.     RKA:
H:       1 rtvgyvydpevlnheckisygatHpenpeRlrlihelLleygllkkmeivtnsqeprkat

             LEEIQLVHSEHHSLLYgtnpldgQKLDPRILLGDDS-Q-KFFSsLPCGGLGVDSDTIWNE
              EE: LVHSE.:  ..        :.L... L   :. . K..        .  ..D..
             deellvHsedyvefl.......eslsktnleelekgtdkill.eielkyfnkgdDtpvf

             LHSSGAARMAVGCVIELASKVASGELKNGFAVVRPPGHHAEESTAMGFCFFNSVAITAKY
             .    AAR:AVG  :ELA.:::.GEL.N.F  .  PGHHA:::.A.GFC:FN:VAI..KY
             aglyeaarlavGgsleladrllegeldnafnwagGPgHHAkkgeasGFCyfNNvAiAiky

             LRDQL--NISKILIVDLDVHHGNGTQQAFYADPSILYISLHRYDEGNFFPGS--GAPNEV
             L .:.   :.::LI:D.DVHHG:GTQ:.FY D..:L :S.H:Y.:G.FFPG:  G  .E:
             llkkyPlyvkRVliiDfDvHHGDGTQeiFydddrVltvSfHkygkGefFPGtieGditei

             GTGLGEGYNINIAWTgglDPPMGDVEYLEAFRTIVKPVAKEFDPDMVLVSAGFDALEGHT
             G.G G:GY.:NI. .   :. ..D .YL.AF: :::PV.::F.PD.:::SAGFDAL G..
             GkgkGkgytlNiPLn...edgtdDesYlsafkhviepvleqFkPdaivisaGfDalygDp

             PPLGGYKVTAKCFGHLTKQLMTLAD----GRVVLALEGGHDLTAICdASEACVNALLGNE
             ..LG.:.:T.: :G.:.: L.:LA.    G :::.LEGG:.L.AI.   ..C  AL.G.
             tqLgsfnLtiegygemvrflkslaqkhcdgpllvvlEGGYtlraia.nvarcwialtggl

             LE .    760
             L
             lg      350
```

Fig. 13

Alignment of SEQ ID NO:2 vs SEQ ID NO:7

```
FASTA (3.34 January 2000) function [optimized, BL50 matrix (15:-5)] ktup: 2
 join: 36, opt: 24, gap-pen: -12/ -2, width:  16
The best scores are:                                      opt
SEQ ID NO:7    Histone Deacetylase                    ( 848) 1067

>> SEQ ID NO:7    Histone Deacetylase                          (848 aa)
 initn: 1063 init1: 1063 opt: 1067
Smith-Waterman score: 1067;  98.125% identity in 160 aa overlap (1-160:518-677)

SEQ ID NO:7 ASEPNLKVRSRLKQKVAERRSSPLLRRKDGNVVTSFKKRMFEVTESSVSSSSPGSGPSSP
                    10        20        30        40        50        60

SEQ ID NO:7 NNGPTGSVTENETSVLPPTPHAEQMVSQQRILIHEDSMNLLSLYTSPSLPNITLGLPAVP
                    70        80        90       100       110       120

SEQ ID NO:7 SQLNASNSLKEKQKCETQTLRQGVPLPGQYGGSIPASSSHPHVTLEGKPPNSSHQALLQH
                   130       140       150       160       170       180

SEQ ID NO:7 LLLKEQMRQQKLLVAGGVPLHPQSPLATKERISPGIRGTHKLPRHRPLNRTQSAPLPQST
                   190       200       210       220       230       240

SEQ ID NO:7 LAQLVIQQQHQQFLEKQKQYQQQIHMNKLLSKSIEQLKQPGSHLEEAEEELQGDQAMQED
                   250       260       270       280       290       300

SEQ ID NO:7 RAPSSGNSTRSDSSACVDDTLGQVGAVKVKEEPVDSDEDAQIQEMESGEQAAFMQQPFLE
                  310       320       330       340       350       360

SEQ ID NO:7 PTHTRALSVRQAPLAAVGMDGLEKHRLVSRTHSSPAASVLPHPAMDRPLQPGSATGIAYD
                   370       380       390       400       410       420

SEQ ID NO:7 PLMLKHQCVCGNSTTHPEHAGRIQSIWSRLQETGLLNKCERIQGRKASLEEIQLVHSEHH
                   430       440       450       460       470       480
```

Fig. 14 (continued)

```
                                                                  10         20
SEQ ID NO:2                                                     VDSDTIWNELHSSGAARMAVGCV
                                                                :::::::::::::::::::::::
SEQ ID NO:7 SLLYGTNPLDGQKLDPRILLGDDSQKFFSSLPCGGLGVDSDTIWNELHSSGAARMAVGCV
                  490        500        510        520        530        540

                30         40         50         60         70         80
SEQ ID NO:2 IELASKVASGELKNGFAVVRPPGHHAEESTAMGFCFFNSVAITAKYLRDQLNISKILIVD
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
SEQ ID NO:7 IELASKVASGELKNGFAVVRPPGHHAEESTAMGFCFFNSVAITAKYLRDQLNISKILIVD
                  550        560        570        580        590        600

                90        100        110        120        130        140
SEQ ID NO:2 LDVHHGNGTQQAFYADPSILYISLHRYDEGNFFPGSGAPNEVGTGLGEGYNINIAWTGGL
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
SEQ ID NO:7 LDVHHGNGTQQAFYADPSILYISLHRYDEGNFFPGSGAPNEVGTGLGEGYNINIAWTGGL
                  610        620        630        640        650        660

              150        160
SEQ ID NO:2 DPPMGDVEYLEAFRLVLLSL
            ::::::::::::::: ..
SEQ ID NO:7 DPPMGDVEYLEAFRTIVKPVAKEFDPDMVLVSAGFDALEGHTPPLGGYKVTAKCFGHLTK
                  670        680        690        700        710        720

SEQ ID NO:7 QLMTLADGRVVLALEGGHDLTAICDASEACVNALLGNELEPLAEDILHQSPNMNAVISLQ
                  730        740        750        760        770        780

SEQ ID NO:7 KIIEIQSKYWKSVRMVAVPRGCALAGAQLQEETETVSALASLTVDVEQPFAQEDSRTAGE
                  790        800        810        820        830        840

SEQ ID NO:7 PMEEEPAL

163 residues in 1 query   sequences; 848 residues in 1 library sequences
```

SEQUENCE LISTING

<110> Bayer AG

<120> REGULATION OF HUMAN HISTONE DEACETYLASE

<130> LIO202 Foreign Countries

<150> US 60/244,183

<151> 2000-10-31

<150> US 60/317,965

<151> 2001-09-10

<160> 7

<170> PatentIn version 3.1

<210> 1

<211> 489

<212> DNA

<213> Homo sapiens

<400> 1

gtggacagtg acaccatttg gaatgagcta cactcgtccg gtgctgcacg catggctgtt 60

ggctgtgtca tcgagctggc ttccaaagtg gcctcaggag agctgaagaa tgggtttgct 120

gttgtgaggc cccctggcca tcacgctgaa gaatccacag ccatggggtt ctgctttttt 180

aattcagttg caattaccgc caaatacttg agagaccaac taaatataag caagatattg 240

attgtagatc tggatgttca ccatggaaac ggtacccagc aggcctttta tgctgacccc 300

agcatcctgt acatttcact ccatcgctat gatgaaggga acttttccc tggcagtgga 360

gccccaaatg aggttggaac aggccttgga gaagggtaca atataaatat tgcctggaca 420

ggtggccttg atcctcccat gggagatgtt gagtaccttg aagcattcag gttggtactt 480

ctttctctc 489

<210> 2

<211> 163

<212> PRT

<213> Homo sapiens

<400> 2

Val Asp Ser Asp Thr Ile Trp Asn Glu Leu His Ser Ser Gly Ala Ala
1 5 10 15

Arg Met Ala Val Gly Cys Val Ile Glu Leu Ala Ser Lys Val Ala Ser
20 25 30

Gly Glu Leu Lys Asn Gly Phe Ala Val Val Arg Pro Pro Gly His His
35 40 45

Ala Glu Glu Ser Thr Ala Met Gly Phe Cys Phe Phe Asn Ser Val Ala
50 55 60

Ile Thr Ala Lys Tyr Leu Arg Asp Gln Leu Asn Ile Ser Lys Ile Leu
65 70 75 80

Ile Val Asp Leu Asp Val His His Gly Asn Gly Thr Gln Gln Ala Phe
85 90 95

Tyr Ala Asp Pro Ser Ile Leu Tyr Ile Ser Leu His Arg Tyr Asp Glu
100 105 110

Gly Asn Phe Phe Pro Gly Ser Gly Ala Pro Asn Glu Val Gly Thr Gly
115 120 125

Leu Gly Glu Gly Tyr Asn Ile Asn Ile Ala Trp Thr Gly Gly Leu Asp
130 135 140

Pro Pro Met Gly Asp Val Glu Tyr Leu Glu Ala Phe Arg Leu Val Leu
145 150 155 160

Leu Ser Leu

<210> 3

<211> 687

<212> PRT

<213> Schizosaccharomyces pombe

<400> 3

Met Leu Ala Ser Asn Ser Asp Gly Ala Ser Thr Ser Val Lys Pro Ser
1 5 10 15

Asp Asp Ala Val Asn Thr Val Thr Pro Trp Ser Ile Leu Leu Thr Asn
20 25 30

Asn Lys Pro Met Ser Gly Ser Glu Asn Thr Leu Asn Asn Glu Ser His
35 40 45

Glu Met Ser Gln Ile Leu Lys Lys Ser Gly Leu Cys Tyr Asp Pro Arg
50 55 60

Met Arg Phe His Ala Thr Leu Ser Glu Val Asp Asp His Pro Glu Asp
65 70 75 80

Pro Arg Arg Val Leu Arg Val Phe Glu Ala Ile Lys Lys Ala Gly Tyr
85 90 95

Val Ser Asn Val Pro Ser Pro Ser Asp Val Phe Leu Arg Ile Pro Ala
100 105 110

Arg Glu Ala Thr Leu Glu Glu Leu Leu Gln Val His Ser Gln Glu Met
115 120 125

Tyr Asp Arg Val Thr Asn Thr Glu Lys Met Ser His Glu Asp Leu Ala
130 135 140

Asn Leu Glu Lys Ile Ser Asp Ser Leu Tyr Tyr Asn Asn Glu Ser Ala
145 150 155 160

```
Phe Cys Ala Arg Leu Ala Cys Gly Ser Ala Ile Glu Thr Cys Thr Ala
                165                 170                 175

Val Val Thr Gly Gln Val Lys Asn Ala Phe Ala Val Val Arg Pro Pro
            180                 185                 190

Gly His His Ala Glu Pro His Lys Pro Gly Gly Phe Cys Leu Phe Asn
        195                 200                 205

Asn Val Ser Val Thr Ala Arg Ser Met Leu Gln Arg Phe Pro Asp Lys
    210                 215                 220

Ile Lys Arg Val Leu Ile Val Asp Trp Asp Ile His His Gly Asn Gly
225                 230                 235                 240

Thr Gln Met Ala Phe Tyr Asp Asp Pro Asn Val Leu Tyr Val Ser Leu
                245                 250                 255

His Arg Tyr Glu Asn Gly Arg Phe Tyr Pro Gly Thr Asn Tyr Gly Cys
            260                 265                 270

Ala Glu Asn Cys Gly Glu Gly Pro Gly Leu Gly Arg Thr Val Asn Ile
        275                 280                 285

Pro Trp Ser Cys Ala Gly Met Gly Asp Gly Asp Tyr Ile Tyr Ala Phe
    290                 295                 300

Gln Arg Val Val Met Pro Val Ala Tyr Glu Phe Asp Pro Asp Leu Val
305                 310                 315                 320

Ile Val Ser Cys Gly Phe Asp Ala Ala Ala Gly Asp His Ile Gly Gln
                325                 330                 335

Phe Leu Leu Thr Pro Ala Ala Tyr Ala His Met Thr Gln Met Leu Met
            340                 345                 350

Gly Leu Ala Asp Gly Lys Val Phe Ile Ser Leu Glu Gly Gly Tyr Asn
        355                 360                 365

Leu Asp Ser Ile Ser Thr Ser Ala Leu Ala Val Ala Gln Ser Leu Leu
    370                 375                 380

Gly Ile Pro Pro Gly Arg Leu His Thr Thr Tyr Ala Cys Pro Gln Ala
385                 390                 395                 400
```

Val Ala Thr Ile Asn His Val Thr Lys Ile Gln Ser Gln Tyr Trp Arg
405 410 415

Cys Met Arg Pro Lys His Phe Asp Ala Asn Pro Lys Asp Ala His Val
420 425 430

Asp Arg Leu His Asp Val Ile Arg Thr Tyr Gln Ala Lys Lys Leu Phe
435 440 445

Glu Asp Trp Lys Ile Thr Asn Met Pro Ile Leu Arg Asp Ser Val Ser
450 455 460

Asn Val Phe Asn Asn Gln Val Leu Cys Ser Ser Asn Phe Phe Gln Lys
465 470 475 480

Asp Asn Leu Leu Val Ile Val His Glu Ser Pro Arg Val Leu Gly Asn
485 490 495

Gly Thr Ser Glu Thr Asn Val Leu Asn Leu Asn Asp Ser Leu Leu Val
500 505 510

Asp Pro Val Ser Leu Tyr Val Glu Trp Ala Met Gln Gln Asp Trp Gly
515 520 525

Leu Ile Asp Ile Asn Ile Pro Glu Val Val Thr Asp Gly Glu Asn Ala
530 535 540

Pro Val Asp Ile Leu Ser Glu Val Lys Glu Leu Cys Leu Tyr Val Trp
545 550 555 560

Asp Asn Tyr Val Glu Leu Ser Ile Ser Lys Asn Ile Phe Phe Ile Gly
565 570 575

Gly Gly Lys Ala Val His Gly Leu Val Asn Leu Ala Ser Ser Arg Asn
580 585 590

Val Ser Asp Arg Val Lys Cys Met Val Asn Phe Leu Gly Thr Glu Pro
595 600 605

Leu Val Gly Leu Lys Thr Ala Ser Glu Glu Asp Leu Pro Thr Trp Tyr
610 615 620

Tyr Arg His Ser Leu Val Phe Val Ser Ser Ser Asn Glu Cys Trp Lys
625 630 635 640

Lys Ala Lys Arg Ala Lys Arg Arg Tyr Gly Arg Leu Met Gln Ser Glu
645 650 655

His Thr Glu Thr Ser Asp Met Met Glu Gln His Tyr Arg Ala Val Thr
660 665 670

Gln Tyr Leu Leu His Leu Leu Gln Lys Ala Arg Pro Thr Ser Gln
675 680 685

<210> 4

<211> 305

<212> DNA

<213> Homo sapiens

<400> 4
aggccttgga gaagggtaca atataaatat tgcctggaca ggtggccttg atcctcccat 60

gggagatgtt gagtaccttg aagcattcag gaccatcgtg aagcctgtgg caaagagttt 120

gatccagaca tggtcttagt atctgctgga tttgatgcat tggaaggcca cacccctcct 180

ctaggagggt acaaagtgac ggcaaaataa actcctgtgc tggaggtaca acagtttgga 240

agtatacttg gggaaagaga aaacacaaga tggaaggaag atctctcttt tcacatcggg 300

agcac 305

<210> 5

<211> 1122

<212> PRT

<213> Homo sapiens

<400> 5

Met Asn Ser Pro Asn Glu Ser Asp Gly Met Ser Gly Arg Glu Pro Ser
1 5 10 15

Leu Glu Ile Leu Pro Arg Thr Ser Leu His Ser Ile Pro Val Thr Val
20 25 30

```
Glu Val Lys Pro Val Leu Pro Arg Ala Met Pro Ser Ser Met Gly Gly
        35                  40                  45

Gly Gly Gly Gly Ser Pro Ser Pro Val Glu Leu Arg Gly Ala Leu Val
    50                  55                  60

Gly Ser Val Asp Pro Thr Leu Arg Glu Gln Gln Leu Gln Gln Glu Leu
65                  70                  75                  80

Leu Ala Leu Lys Gln Gln Gln Gln Leu Gln Lys Gln Leu Leu Phe Ala
                85                  90                  95

Glu Phe Gln Lys Gln His Asp His Leu Thr Arg Gln His Glu Val Gln
            100                 105                 110

Leu Gln Lys His Leu Lys Gln Gln Gln Glu Met Leu Ala Ala Lys Gln
        115                 120                 125

Gln Gln Glu Met Leu Ala Ala Lys Arg Gln Gln Glu Leu Glu Gln Gln
    130                 135                 140

Arg Gln Arg Glu Gln Gln Arg Gln Glu Glu Leu Glu Lys Gln Arg Leu
145                 150                 155                 160

Glu Gln Gln Leu Leu Ile Leu Arg Asn Lys Glu Lys Ser Lys Glu Ser
                165                 170                 175

Ala Ile Ala Ser Thr Glu Val Lys Leu Arg Leu Gln Glu Phe Leu Leu
            180                 185                 190

Ser Lys Ser Lys Glu Pro Thr Pro Gly Gly Leu Asn His Ser Leu Pro
        195                 200                 205

Gln His Pro Lys Cys Trp Gly Ala His His Ala Ser Leu Asp Gln Ser
    210                 215                 220

Ser Pro Pro Gln Ser Gly Pro Pro Gly Thr Pro Pro Ser Tyr Lys Leu
225                 230                 235                 240

Pro Leu Pro Gly Pro Tyr Asp Ser Arg Asp Asp Phe Pro Leu Arg Lys
                245                 250                 255
```

Thr Ala Ser Glu Pro Asn Leu Lys Val Arg Ser Arg Leu Lys Gln Lys
260 265 270

Val Ala Glu Arg Arg Ser Ser Pro Leu Leu Arg Arg Lys Asp Gly Thr
275 280 285

Val Ile Ser Thr Phe Lys Lys Arg Ala Val Glu Ile Thr Gly Ala Gly
290 295 300

Pro Gly Ala Ser Ser Val Cys Asn Ser Ala Pro Gly Ser Gly Pro Ser
305 310 315 320

Ser Pro Asn Ser Ser His Ser Thr Ile Ala Glu Asn Gly Phe Thr Gly
325 330 335

Ser Val Pro Asn Ile Pro Thr Glu Met Leu Pro Gln His Arg Ala Leu
340 345 350

Pro Leu Asp Ser Ser Pro Asn Gln Phe Ser Leu Tyr Thr Ser Pro Ser
355 360 365

Leu Pro Asn Ile Ser Leu Gly Leu Gln Ala Thr Val Thr Val Thr Asn
370 375 380

Ser His Leu Thr Ala Ser Pro Lys Leu Ser Thr Gln Gln Glu Ala Glu
385 390 395 400

Arg Gln Ala Leu Gln Ser Leu Arg Gln Gly Gly Thr Leu Thr Gly Lys
405 410 415

Phe Met Ser Thr Ser Ser Ile Pro Gly Cys Leu Leu Gly Val Ala Leu
420 425 430

Glu Gly Asp Gly Ser Pro His Gly His Ala Ser Leu Leu Gln His Val
435 440 445

Leu Leu Leu Glu Gln Ala Arg Gln Gln Ser Thr Leu Ile Ala Val Pro
450 455 460

Leu His Gly Gln Ser Pro Leu Val Thr Gly Glu Arg Val Ala Thr Ser
465 470 475 480

Met Arg Thr Val Gly Lys Leu Pro Arg His Arg Pro Leu Ser Arg Thr
485 490 495

Gln Ser Ser Pro Leu Pro Gln Ser Pro Gln Ala Leu Gln Gln Leu Val
500 505 510

Met Gln Gln Gln His Gln Gln Phe Leu Glu Lys Gln Lys Gln Gln Gln
515 520 525

Leu Gln Leu Gly Lys Ile Leu Thr Lys Thr Gly Glu Leu Pro Arg Gln
530 535 540

Pro Thr Thr His Pro Glu Glu Thr Glu Glu Glu Leu Thr Glu Gln Gln
545 550 555 560

Glu Val Leu Leu Gly Glu Gly Ala Leu Thr Met Pro Arg Glu Gly Ser
565 570 575

Thr Glu Ser Glu Ser Thr Gln Glu Asp Leu Glu Glu Glu Asp Glu Glu
580 585 590

Glu Asp Gly Glu Glu Glu Glu Asp Cys Ile Gln Val Lys Asp Glu Glu
595 600 605

Gly Glu Ser Gly Ala Glu Glu Gly Pro Asp Leu Glu Glu Pro Gly Ala
610 615 620

Gly Tyr Lys Lys Leu Phe Ser Asp Ala Gln Pro Leu Gln Pro Leu Gln
625 630 635 640

Val Tyr Gln Ala Pro Leu Ser Leu Ala Thr Val Pro His Gln Ala Leu
645 650 655

Gly Arg Thr Gln Ser Ser Pro Ala Ala Pro Gly Gly Met Lys Ser Pro
660 665 670

Pro Asp Gln Pro Val Lys His Leu Phe Thr Thr Gly Val Val Tyr Asp
675 680 685

Thr Phe Met Leu Lys His Gln Cys Met Cys Gly Asn Thr His Val His
690 695 700

Pro Glu His Ala Gly Arg Ile Gln Ser Ile Trp Ser Arg Leu Gln Glu
705 710 715 720

Thr Gly Leu Leu Ser Lys Cys Glu Arg Ile Arg Gly Arg Lys Ala Thr
725 730 735

Leu Asp Glu Ile Gln Thr Val His Ser Glu Tyr His Thr Leu Leu Tyr
740 745 750

Gly Thr Ser Pro Leu Asn Arg Gln Lys Leu Asp Ser Lys Lys Leu Leu
755 760 765

Gly Pro Ile Ser Gln Lys Met Tyr Ala Val Leu Pro Cys Gly Gly Ile
770 775 780

Gly Val Asp Ser Asp Thr Val Trp Asn Glu Met His Ser Ser Ser Ala
785 790 795 800

Val Arg Met Ala Val Gly Cys Leu Leu Glu Leu Ala Phe Lys Val Ala
805 810 815

Ala Gly Glu Leu Lys Asn Gly Phe Ala Ile Ile Arg Pro Pro Gly His
820 825 830

His Ala Glu Glu Ser Thr Ala Met Gly Phe Cys Phe Phe Asn Ser Val
835 840 845

Ala Ile Thr Ala Lys Leu Leu Gln Gln Lys Leu Asn Val Gly Lys Val
850 855 860

Leu Ile Val Asp Trp Asp Ile His His Gly Asn Gly Thr Gln Gln Ala
865 870 875 880

Phe Tyr Asn Asp Pro Ser Val Leu Tyr Ile Ser Leu His Arg Tyr Asp
885 890 895

Asn Gly Asn Phe Phe Pro Gly Ser Gly Ala Pro Glu Glu Val Gly Gly
900 905 910

Gly Pro Gly Val Gly Tyr Asn Val Asn Val Ala Trp Thr Gly Gly Val
915 920 925

Asp Pro Pro Ile Gly Asp Val Glu Tyr Leu Thr Ala Phe Arg Thr Val
930 935 940

Val Met Pro Ile Ala His Glu Phe Ser Pro Asp Val Val Leu Val Ser
945 950 955 960

Ala Gly Phe Asp Ala Val Glu Gly His Leu Ser Pro Leu Gly Gly Tyr
965 970 975

Ser Val Thr Ala Arg Cys Phe Gly His Leu Thr Arg Gln Leu Met Thr
980 985 990

Leu Ala Gly Gly Arg Val Val Leu Ala Leu Glu Gly Gly His Asp Leu
995 1000 1005

Thr Ala Ile Cys Asp Ala Ser Glu Ala Cys Val Ser Ala Leu Leu
1010 1015 1020

Ser Val Glu Leu Gln Pro Leu Asp Glu Ala Val Leu Gln Gln Lys
1025 1030 1035

Pro Asn Ile Asn Ala Val Ala Thr Leu Glu Lys Val Ile Glu Ile
1040 1045 1050

Gln Ser Lys His Trp Ser Cys Val Gln Lys Phe Ala Ala Gly Leu
1055 1060 1065

Gly Arg Ser Leu Arg Glu Ala Gln Ala Gly Glu Thr Glu Glu Ala
1070 1075 1080

Glu Thr Val Ser Ala Met Ala Leu Leu Ser Val Gly Ala Glu Gln
1085 1090 1095

Ala Gln Ala Ala Ala Ala Arg Glu His Ser Pro Arg Pro Ala Glu
1100 1105 1110

Glu Pro Met Glu Gln Glu Pro Ala Leu
1115 1120

<210> 6

<211> 2544

<212> DNA

<213> Homo sapiens

<400> 6

gcctctgagc ccaacttgaa ggtgcggtcc aggttaaaac agaaagtggc agagaggaga 60

agcagcccct tactcaggcg gaaggatgga aatgttgtca cttcattcaa gaagcgaatg 120

tttgaggtga cagaatcctc agtcagtagc agttctccag gctctggtcc cagttcacca 180

aacaatgggc caactggaag tgttactgaa aatgagactt cggttttgcc ccctaccccc 240

catgccgagc aaatggtttc acagcaacgc attctaattc atgaagattc catgaacctg 300

ctaagtcttt atacctctcc ttctttgccc aacattacct tggggcttcc cgcagtgcca 360

tcccagctca atgcttcgaa ttcactcaaa gaaaagcaga agtgtgagac gcagacgctt 420

aggcaaggtg ttcctctgcc tgggcagtat ggaggcagca tcccggcatc ttccagccac 480

cctcatgtta ctttagaggg aaagccaccc aacagcagcc accaggctct cctgcagcat 540

ttattattga aagaacaaat gcgacagcaa aagcttcttg tagctggtgg agttcccttа 600

catcctcagt ctcccttggc aacaaaagag agaatttcac ctggcattag aggtacccac 660

aaattgcccc gtcacagacc cctgaaccga acccagtctg caccttтgcc tcagagcacg 720

ttggctcagc tggtcattca acagcaacac cagcaattct tggagaagca gaagcaatac 780

cagcagcaga tccacatgaa caaactgctt tcgaaatcta ttgaacaact gaagcaacca 840

ggcagtcacc ttgaggaagc agaggaagag cttcaggggg accaggcgat gcaggaagac 900

agagcgccct ctagtggcaa cagcactagg agcgacagca gtgcttgtgt ggatgacaca 960

ctgggacaag ttggggctgt gaaggtcaag gaggaaccag tggacagtga tgaagatgct 1020

cagatccagg aaatggaatc tggggagcag gctgctttta tgcaacagcc tttcctggaa 1080

cccacgcaca cacgtgcgct ctctgtgcgc caagctccgc tggctgcggt tggcatggat 1140

ggattagaga aacaccgtct cgtctccagg actcactctt cccctgctgc ctctgtttta 1200

cctcacccag caatggaccg ccccctccag cctggctctg caactggaat tgcctatgac 1260

cccttgatgc tgaaacacca gtgcgtttgt ggcaattcca ccacccaccc tgagcatgct 1320

ggacgaatac agagtatctg gtcacgactg caagaaactg ggctgctaaa taaatgtgag 1380

cgaattcaag gtcgaaaagc cagcctggag gaaatacagc ttgttcattc tgaacatcac 1440

tcactgttgt atggcaccaa ccccctggac ggacagaagc tggaccccag gatactccta 1500

ggtgatgact ctcaaaagtt tttttcctca ttaccttgtg gtggacttgg ggtggacagt 1560

gacaccattt ggaatgagct acactcgtcc ggtgctgcac gcatggctgt tggctgtgtc 1620

atcgagctgg cttccaaagt ggcctcagga gagctgaaga atgggtttgc tgttgtgagg 1680

ccccctggcc atcacgctga agaatccaca gccatggggt tctgcttttt taattcagtt 1740

gcaattaccg ccaaatactt gagagaccaa ctaaatataa gcaagatatt gattgtagat 1800

ctggatgttc accatggaaa cggtacccag caggcctttt atgctgaccc cagcatcctg 1860

tacatttcac tccatcgcta tgatgaaggg aacttttcc ctggcagtgg agccccaaat 1920

gaggttggaa caggccttgg agaagggtac aatataaata ttgcctggac aggtggcctt 1980

gatcctccca tgggagatgt tgagtacctt gaagcattca ggaccatcgt gaagcctgtg 2040

gccaaagagt ttgatccaga catggtctta gtatctgctg gatttgatgc attggaaggc 2100

cacacccctc ctctaggagg gtacaaagtg acggcaaaat gttttggtca tttgacgaag 2160

caattgatga cattggctga tggacgtgtg gtgttggctc tagaaggagg acatgatctc 2220

acagccatct gtgatgcatc agaagcctgt gtaaatgccc ttctaggaaa tgagctggag 2280

ccacttgcag aagatattct ccaccaaagc ccgaatatga atgctgttat ttctttacag 2340

aagatcattg aaattcaaag caagtattgg aagtcagtaa ggatggtggc tgtgccaagg 2400

ggctgtgctc tggctggtgc tcagttgcaa gaggagacag agaccgtttc tgccctggcc 2460

tccctaacag tggatgtgga acagcccttt gctcaggaag acagcagaac tgctggtgag 2520

cctatggaag aggagccagc cttg 2544

<210> 7

<211> 848

<212> PRT

<213> Homo sapiens

<400> 7

Ala Ser Glu Pro Asn Leu Lys Val Arg Ser Arg Leu Lys Gln Lys Val
1 5 10 15

Ala Glu Arg Arg Ser Ser Pro Leu Leu Arg Arg Lys Asp Gly Asn Val
20 25 30

Val Thr Ser Phe Lys Lys Arg Met Phe Glu Val Thr Glu Ser Ser Val
35 40 45

Ser Ser Ser Ser Pro Gly Ser Gly Pro Ser Ser Pro Asn Asn Gly Pro
50 55 60

```
Thr Gly Ser Val Thr Glu Asn Glu Thr Ser Val Leu Pro Pro Thr Pro
65                  70                  75                  80

His Ala Glu Gln Met Val Ser Gln Gln Arg Ile Leu Ile His Glu Asp
                85                  90                  95

Ser Met Asn Leu Leu Ser Leu Tyr Thr Ser Pro Ser Leu Pro Asn Ile
            100                 105                 110

Thr Leu Gly Leu Pro Ala Val Pro Ser Gln Leu Asn Ala Ser Asn Ser
        115                 120                 125

Leu Lys Glu Lys Gln Lys Cys Glu Thr Gln Thr Leu Arg Gln Gly Val
    130                 135                 140

Pro Leu Pro Gly Gln Tyr Gly Gly Ser Ile Pro Ala Ser Ser Ser His
145                 150                 155                 160

Pro His Val Thr Leu Glu Gly Lys Pro Pro Asn Ser Ser His Gln Ala
                165                 170                 175

Leu Leu Gln His Leu Leu Leu Lys Glu Gln Met Arg Gln Gln Lys Leu
            180                 185                 190

Leu Val Ala Gly Gly Val Pro Leu His Pro Gln Ser Pro Leu Ala Thr
        195                 200                 205

Lys Glu Arg Ile Ser Pro Gly Ile Arg Gly Thr His Lys Leu Pro Arg
    210                 215                 220

His Arg Pro Leu Asn Arg Thr Gln Ser Ala Pro Leu Pro Gln Ser Thr
225                 230                 235                 240

Leu Ala Gln Leu Val Ile Gln Gln Gln His Gln Gln Phe Leu Glu Lys
                245                 250                 255

Gln Lys Gln Tyr Gln Gln Gln Ile His Met Asn Lys Leu Leu Ser Lys
            260                 265                 270

Ser Ile Glu Gln Leu Lys Gln Pro Gly Ser His Leu Glu Glu Ala Glu
        275                 280                 285

Glu Glu Leu Gln Gly Asp Gln Ala Met Gln Glu Asp Arg Ala Pro Ser
    290                 295                 300
```

```
Ser Gly Asn Ser Thr Arg Ser Asp Ser Ser Ala Cys Val Asp Asp Thr
305                 310                 315                 320

Leu Gly Gln Val Gly Ala Val Lys Val Lys Glu Glu Pro Val Asp Ser
                325                 330                 335

Asp Glu Asp Ala Gln Ile Gln Glu Met Glu Ser Gly Glu Gln Ala Ala
            340                 345                 350

Phe Met Gln Gln Pro Phe Leu Glu Pro Thr His Thr Arg Ala Leu Ser
        355                 360                 365

Val Arg Gln Ala Pro Leu Ala Ala Val Gly Met Asp Gly Leu Glu Lys
    370                 375                 380

His Arg Leu Val Ser Arg Thr His Ser Ser Pro Ala Ala Ser Val Leu
385                 390                 395                 400

Pro His Pro Ala Met Asp Arg Pro Leu Gln Pro Gly Ser Ala Thr Gly
                405                 410                 415

Ile Ala Tyr Asp Pro Leu Met Leu Lys His Gln Cys Val Cys Gly Asn
            420                 425                 430

Ser Thr Thr His Pro Glu His Ala Gly Arg Ile Gln Ser Ile Trp Ser
        435                 440                 445

Arg Leu Gln Glu Thr Gly Leu Leu Asn Lys Cys Glu Arg Ile Gln Gly
    450                 455                 460

Arg Lys Ala Ser Leu Glu Glu Ile Gln Leu Val His Ser Glu His His
465                 470                 475                 480

Ser Leu Leu Tyr Gly Thr Asn Pro Leu Asp Gly Gln Lys Leu Asp Pro
                485                 490                 495

Arg Ile Leu Leu Gly Asp Asp Ser Gln Lys Phe Phe Ser Ser Leu Pro
            500                 505                 510

Cys Gly Gly Leu Gly Val Asp Ser Asp Thr Ile Trp Asn Glu Leu His
        515                 520                 525
```

```
Ser Ser Gly Ala Ala Arg Met Ala Val Gly Cys Val Ile Glu Leu Ala
    530                 535                 540

Ser Lys Val Ala Ser Gly Glu Leu Lys Asn Gly Phe Ala Val Val Arg
545                 550                 555                 560

Pro Pro Gly His His Ala Glu Glu Ser Thr Ala Met Gly Phe Cys Phe
                565                 570                 575

Phe Asn Ser Val Ala Ile Thr Ala Lys Tyr Leu Arg Asp Gln Leu Asn
            580                 585                 590

Ile Ser Lys Ile Leu Ile Val Asp Leu Asp Val His His Gly Asn Gly
        595                 600                 605

Thr Gln Gln Ala Phe Tyr Ala Asp Pro Ser Ile Leu Tyr Ile Ser Leu
    610                 615                 620

His Arg Tyr Asp Glu Gly Asn Phe Phe Pro Gly Ser Gly Ala Pro Asn
625                 630                 635                 640

Glu Val Gly Thr Gly Leu Gly Glu Gly Tyr Asn Ile Asn Ile Ala Trp
                645                 650                 655

Thr Gly Gly Leu Asp Pro Pro Met Gly Asp Val Glu Tyr Leu Glu Ala
            660                 665                 670

Phe Arg Thr Ile Val Lys Pro Val Ala Lys Glu Phe Asp Pro Asp Met
        675                 680                 685

Val Leu Val Ser Ala Gly Phe Asp Ala Leu Glu Gly His Thr Pro Pro
    690                 695                 700

Leu Gly Gly Tyr Lys Val Thr Ala Lys Cys Phe Gly His Leu Thr Lys
705                 710                 715                 720

Gln Leu Met Thr Leu Ala Asp Gly Arg Val Val Leu Ala Leu Glu Gly
                725                 730                 735

Gly His Asp Leu Thr Ala Ile Cys Asp Ala Ser Glu Ala Cys Val Asn
            740                 745                 750

Ala Leu Leu Gly Asn Glu Leu Glu Pro Leu Ala Glu Asp Ile Leu His
        755                 760                 765
```

```
Gln Ser Pro Asn Met Asn Ala Val Ile Ser Leu Gln Lys Ile Ile Glu
    770                 775                 780

Ile Gln Ser Lys Tyr Trp Lys Ser Val Arg Met Val Ala Val Pro Arg
785                 790                 795                 800

Gly Cys Ala Leu Ala Gly Ala Gln Leu Gln Glu Glu Thr Glu Thr Val
                805                 810                 815

Ser Ala Leu Ala Ser Leu Thr Val Asp Val Glu Gln Pro Phe Ala Gln
            820                 825                 830

Glu Asp Ser Arg Thr Ala Gly Glu Pro Met Glu Glu Glu Pro Ala Leu
        835                 840                 845
```

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 May 2002 (10.05.2002)

PCT

(10) International Publication Number
WO 02/036783 A3

(51) International Patent Classification[7]: C12N 15/55, 9/12, C12Q 1/68, 1/44, A61K 31/713, 38/46, C12N 15/62, C07K 16/40, C12N 15/11, 9/16 // A61P 35/00

(21) International Application Number: PCT/EP01/12517

(22) International Filing Date: 30 October 2001 (30.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/244,183	31 October 2000 (31.10.2000)	US
60/317,965	10 September 2001 (10.09.2001)	US

(71) Applicant *(for all designated States except US)*: **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **XIAO, Yonghong** [US/US]; 75 Dana Street, #1, Cambridge, MA 02138 (US).

(74) Common Representative: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
1 May 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/036783 A3

(54) Title: REGULATION OF HUMAN HISTONE DEACETYLASE

(57) Abstract: Reagents which regulate human histone deacetylase and reagents which bind to human histone deacetylase gene products can play a role in preventing, ameliorating, or correcting dysfunctions or diseases including, but not limited to, cancer.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 01/12517

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/55 C12N9/12 C12Q1/68 C12Q1/44 A61K31/713
A61K38/46 C12N15/62 C07K16/40 C12N15/11 C12N9/16
//A61P35/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
BIOSIS, EPO-Internal, EMBL, BIOTECHNOLOGY ABS, CHEM ABS Data, EMBASE, LIFESCIENCES, MEDLINE, PAJ, SCISEARCH, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WANG AUDREY H ET AL: "HDAC4, a human histone deacetylase related to yeast HDA1, is a transcriptional corepressor." MOLECULAR AND CELLULAR BIOLOGY, vol. 19, no. 11, November 1999 (1999-11), pages 7816-7827, XP002220348 ISSN: 0270-7306 figure 1	1-17,32, 35-48, 60-65, 68-71
X	-& DATABASE EMBL 'Online! EMBL; Homo sapiens histone deacetylase HDAC7, 27 October 1999 (1999-10-27) WANG A.H. ET AL.: "HDAC4, a human histone deacetylase related to yeast HDA1, is a transcriptional corepressor" Database accession no. AF124924 XP002220350 abstract	1-17,32, 35-48, 60-65, 68-71

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 November 2002	26/11/2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Pilat, D

1

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GROZINGER CHRISTINA M ET AL: "Three proteins define a class of human histone deacetylases related to yeast Hda1p" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 96, no. 9, 27 April 1999 (1999-04-27), pages 4868-4873, XP002159644 ISSN: 0027-8424 the whole document	1-17,32, 35-48, 60-65, 68-71
X	-& DATABASE EMBL 'Online! EMBL; Homo sapiens histone deacetylase 4 mRNA, 6 May 1999 (1999-05-06) GROZINGER C.M. ET AL: "Three proteins define a class of human histone deacetylases related to yeast Hda1p" Database accession no. AF132607 XP002220351 abstract	1-17,32, 35-48, 60-65, 68-71
X	-& DATABASE EMBL 'Online! EMBL; Homo sapiens histone deacetylase 5 mRNA, 6 May 1999 (1999-05-06) "Three proteins define a class of human histone deacetylases related to yeast Hdap1" Database accession no. AF132608 XP002220352 abstract	1-17,32, 35-48, 60-65, 68-71
X	NAGASE T ET AL: "PREDICTION OF THE CODING SEQUENCES OF UNIDENTIFIED HUMAN GENES. XI. THE COMPLETE SEQUENCES OF 100 NEW CDNA CLONES FROM BRAIN WHICH CODE FOR LARGE PROTEINS IN VITRO" DNA RESEARCH, UNIVERSAL ACADEMY PRESS, JP, vol. 5, 1998, pages 277-286, XP000828191 ISSN: 1340-2838	1-17,32, 35-48, 60-65, 68-71
X	-& DATABASE EMBL 'Online! EMBL; Homo sapiens mRNA for KIAA0744, 17 November 1998 (1998-11-17) NAGASE T. ET AL.: "Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro." Database accession no. AB018287 XP002220353 abstract	1-17,32, 35-48, 60-65, 68-71

-/--

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	ZHOU XIANBO ET AL: "Cloning and characterization of a histone deacetylase, HDAC9." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 98, no. 19, 11 September 2001 (2001-09-11), pages 10572-10577, XP002220349 September 11, 2001 ISSN: 0027-8424 the whole document	1-17,32, 35-48, 60-65, 68-71
P,X	-& DATABASE EMBL 'Online! EMBL; Homo sapiens histone deacetylase 9a (HDAC9) mRNA, 13 September 2001 (2001-09-13) ZHOU X. ET AL.: "Cloning and characterization of a histone deacetylase, HDAC9" Database accession no. AY032738 XP002220354 abstract	1-17,32, 35-48, 60-65, 68-71
P,X	-& DATABASE EMBL 'Online! EMBL; Homo sapiens histone deacetylase 9 (HDAC9) mRNA, 13 September 2001 (2001-09-13) ZHOU: ""Cloning and characterization of a histone deacetylase, HDAC9"" Database accession no. AY032737 XP002220355 abstract	1-17,32, 35-48, 60-65, 68-71
P,X	WO 00 71703 A (METHYLGENE INC) 30 November 2000 (2000-11-30) SEQ ID N°33 page 59 -page 62; examples 5,6 page 64 -page 65	8-17, 36-48, 62,64, 68-71

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP 01/12517

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 68-71 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 8,9,13-17,33,34, 52-65, 68-71 partially
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 8,9,13-17,33,34, 52-65, 68-71 partially

Present claims 8,13,33,52 relate to a method defined by reference to a desirable characteristic or property, namely contacting a biological sample with a reagent which specifcally interacts with the polynucleotide or the histone deacetylase polypeptide

The claims cover all methods having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such methods. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the method by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to the methods which comprises the steps of contacting a biological sample with a nucleotide as claimed in claim 1 a),b), or an antibody specifically binding to the polypeptide of claim 4 or an antisense oligonucleotide or a ribozyme specifically binding to SEQ ID N°1 or 6.
Claims 9, 34, 53-59 which refer directly or indirectly to claims 8, 33 and 52 were restricted accordingly.

Present claims 14, 60 and 62 relates to a product defined by reference to a desirable characteristic or property, namely that modulates the activity of a histone deacetylase polypeptide or the polynucleotide encoding a histone deacetylase. The claims cover all products having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such products. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to antibodies specifically binding the polypeptide of SEQ ID N°2 or 7, to antisense oligonucleotides comprising 11 contiguous nucleotide of SEQ ID N° 1 or 6, ribozyme specifically binding to the mRNA transcribed from SEQ ID N°1 or 6 have been searched (see p.28 line 4, p.31 line 15, p.33 line 33).
Insofar as claims 15-17, 61 and 63-65 refers directly or indirectly to claims 14, 60 and 62 the search was restricted accordingly.
The search has been carried out and based on the alleged effects of the compound/composition for claims 68-71 but was, in line with the clarity

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

objection raised, limited to those reagents as described hereinabove.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 01/12517

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0071703 A	30-11-2000	AU 6718200 A	12-12-2000
		EP 1173562 A2	23-01-2002
		WO 0071703 A2	30-11-2000